ANALYSIS & PDE

Volume 12 No. 5 2019

XUAN THINH DUONG, JI LI, YUMENG OU,
JILL PIPHER AND BRETT D. WICK

COMMUTATORS OF MULTIPARAMETER FLAG SINGULAR INTEGRALS AND APPLICATIONS

msp

ANALYSIS AND PDE
Vol. 12, No. 5, 2019
dx.doi.org/10.2140/apde.2019.12.1325

COMMUTATORS OF MULTIPARAMETER FLAG SINGULAR INTEGRALS AND APPLICATIONS

XUAN THINH DUONG, JI LI, YUMENG OU, JILL PIPHER AND BRETT D. WICK

We introduce the iterated commutator for the Riesz transforms in the multiparameter flag setting, and prove the upper bound of this commutator with respect to the symbol b in the flag BMO space. Our methods require the techniques of semigroups, harmonic functions and multiparameter flag Littlewood–Paley analysis. We also introduce the *big* commutator in this multiparameter flag setting and prove the upper bound with symbol b in the flag little bmo space by establishing the "exponential-logarithmic" bridge between this flag little bmo space and the Muckenhoupt A_p weights with flag structure. As an application, we establish the div-curl lemmas with respect to the appropriate Hardy spaces in the multiparameter flag setting.

1. Introduction and statement of main results

The Calderón–Zygmund theory of singular integrals has been central to the success and applicability of modern harmonic analysis in the last fifty years. This theory has had extensive applications to other fields of mathematics such as complex analysis, geometric measure theory and partial differential equations. In the setting of Euclidean spaces $\mathbb{R}^n$, a notable property of standard Calderón–Zygmund singular integrals, shared with the Hardy–Littlewood maximal operator, is that these operators commute with the classical one-parameter family of dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta x_1, \dots, \delta x_n)$ for $\delta > 0$. See for example [Stein 1993].

The product Calderón–Zygmund theory in harmonic analysis was introduced in the 1970s and has been studied extensively since then. The model case is a tensor product of classical singular integral operators; such operators arise in the context of questions about summation of multiple variable Fourier series. Early key work in this field includes that of Chang and R. Fefferman [1980; 1982; 1985], R. Fefferman [1986; 1987; 1999], R. Fefferman and Stein [1982], C. Fefferman and Stein [1972], Gundy and Stein [1979], Journé [1985], and Pipher [1986]. Included in these works are the identification of appropriate notions of product BMO space and product Hardy space $H^p(\mathbb{R}^n \times \mathbb{R}^m)$.

More recently, the theory of (iterated) commutators has been developed in connection with the Chang–Fefferman BMO space, including paraproducts and multiparameter div-curl lemmas; see, for example, [Dalenc and Ou 2016; Ferguson and Lacey 2002; Ferguson and Sadosky 2000; Lacey et al. 2009; 2010; 2012; Lacey and Terwilleger 2009]. In contrast with the classical Euclidean setting, the product Calderón–Zygmund singular integrals and the *strong* maximal function operator commute with the multiparameter dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta_1 x_1, \dots, \delta_n x_n)$ for $\delta = (\delta_1, \dots, \delta_n) \in (0, \infty)^n$.

MSC2010: 42B30, 42B20, 42B35.
Keywords: multiparameter flag setting, flag commutator, Hardy space, BMO space, div-curl lemma .

A new type of multiparameter structure, which lies in between one-parameter and tensor product, was introduced by Müller, Ricci and Stein in [Müller et al. 1995; 1996], where they studied the L^p boundedness of Marcinkiewicz multipliers $m(\mathcal{L}, iT)$ on the Heisenberg group, where $\mathcal{L}$ is the sub-Laplacian and T is the central invariant vector field, with m being a multiplier of Marcinkiewicz-type. They showed that such Marcinkiewicz multipliers can be characterized by a convolution operator $f * K$, where K is a so-called *flag* convolution kernel. This multiparameter flag structure is not explicit, but only *implicit* in the sense that one cannot formulate it in terms of an explicit dilation δ acting on x. Later, the notion of flag kernels (having singularities on appropriate flag varieties) and the properties of the corresponding singular integrals were then extended to the higher-step case by Nagel, Ricci and Stein [Nagel et al. 2001] on Euclidean space and their applications on certain quadratic CR submanifolds of $\mathbb{C}^n$. Recently, Nagel, Ricci, Stein and Wainger [Nagel et al. 2012; 2018] established the theory of singular integrals with flag kernels in a more general setting of homogeneous groups. They proved that, on a homogeneous group, singular integral operators with flag kernels are bounded on L^p, $1 < p < \infty$, and form an algebra. (See also [Głowacki 2010] for related work.) Associated to this implicit multiparameter flag structure, the Hardy space $H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ and BMO space $\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ were introduced by Han, Lu and Sawyer [Han and Lu 2008; Han et al. 2014] through their creation of a flag-type Littlewood–Paley theory. More recently, Han, Lee, and the second and fifth authors [Han et al. 2016a] established a full characterization of $H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ via appropriate flag-type nontangential, radial maximal functions, Littlewood–Paley theory via Poisson integrals, the flag-type Riesz transforms, as well as flag atomic decompositions.

In the multiparameter setting, the dilation structure $\delta \cdot x = (\delta_1 x_1, \ldots, \delta_n x_n)$ for $\delta := (\delta_1, \ldots, \delta_n) \in (0, \infty)^n$ determines a geometry that is reflected by axes-parallel rectangles of arbitrary side-lengths. Indeed, the strong maximal function is defined as the supremum of averages over such rectangles, and the Chang–Fefferman product BMO space can also be characterized using such rectangles. When it comes to the flag setting, the lack of an explicit dilation structure makes its geometry much more obscure. However, from the study of properties of the flag singular integrals, such as the flag Riesz transforms that will be introduced below, one realizes that the flag geometry can be reflected by axes-parallel rectangles with certain restriction on the side-lengths. For example, the flag rectangles in $\mathbb{R}^n \times \mathbb{R}^m$ are the ones of the form $R = I \times J \subset \mathbb{R}^n \times \mathbb{R}^m$ with $\ell(I) \le \ell(J)$. Compared to the multiparameter setting, the restriction $\ell(I) \le \ell(J)$ gives rise to new difficulties. For instance, a very useful trick in the study of problems in the multiparameter setting is to take a sequence of rectangles $\{I \times J_i\}$ and let J_i shrink to a point y_0 as $i \to \infty$. This can usually effectively reduce the problem to one-parameter. However, in the flag setting, such an operation is not allowed anymore. Other intrinsic difficulties of the flag setting can be better described from the analytic perspective, which will be discussed below.

A commutator of a classical Calderón–Zygmund singular integral with a BMO function is a bounded operator on L^p with norm equivalent to the BMO norm of the symbol [Coifman et al. 1976]. Modern methods of proving the upper bound of these commutators in the multiparameter product setting rely upon the existence of a wavelet basis for $L^2(\mathbb{R}^n)$, such as the Meyer wavelets or Haar wavelets; see for example [Lacey et al. 2009; Dalenc and Ou 2016]. It turns out that the behavior of the commutator is straightforward to analyze in terms of the wavelet basis. One method of proof shows that the commutator can be written as a linear combination of paraproducts and simple wavelet analogs of the Calderón–Zygmund operator in question. The other approach uses the wavelet basis to dominate the commutator by a composition of sparse operators. In the flag setting, we lack a suitable wavelet basis and this approach is not available. Essentially, the wavelet basis requires the construction of a suitable multiresolution analysis, which we do not have in this flag setting. Hence, instead of the wavelet basis, we resort to using a method based on heat semigroups and flag-type Littlewood–Paley theory, exploiting the connection between the Reisz transforms and the Laplacian.

We now recall the flag Riesz transforms as studied in [Han et al. 2016a]. We use $R_j^{(1)}$ to denote the j-th Riesz transform on $\mathbb{R}^{n+m}$, $j = 1, 2, \dots, n+m$, and we use $R_k^{(2)}$ to denote the k-th Riesz transform on $\mathbb{R}^m$, $k = 1, 2, \dots, m$. Namely, we have that for $g^{(1)} \in L^2(\mathbb{R}^{n+m})$,

$$R_j^{(1)} g^{(1)}(x) = \text{p.v. } c_{n+m} \int_{\mathbb{R}^{n+m}} \frac{x_j - y_j}{|x-y|^{n+m+1}} g^{(1)}(y)\, dy, \quad x \in \mathbb{R}^{n+m},$$

and for $g^{(2)} \in L^2(\mathbb{R}^m)$,

$$R_k^{(2)} g^{(2)}(z) = \text{p.v. } c_m \int_{\mathbb{R}^m} \frac{w_j - z_j}{|w-z|^{m+1}} g^{(2)}(w)\, dw, \quad z \in \mathbb{R}^m.$$

For $f \in L^2(\mathbb{R}^{n+m})$, we set

$$R_{j,k}(f) = R_j^{(1)} * R_k^{(2)} *_2 f; \tag{1-1}$$

that is, $R_{j,k}$ is the composition of $R_j^{(1)}$ and $R_k^{(2)}$. Note that the flag structure appears in $R_{j,k}$.

Given two functions $b, f \in L^2(\mathbb{R}^{n+m})$, we first recall the usual definition of commutator

$$[b, R_j^{(1)}](f)(x_1, x_2) := b(x_1, x_2) R_j^{(1)} * f(x_1, x_2) - R_j^{(1)} * (bf)(x_1, x_2). \tag{1-2}$$

The commutator can also act only on the second variable:

$$[b, R_k^{(2)}]_2(f)(x_1, x_2) := b(x_1, x_2) R_k^{(2)} *_2 f(x_1, x_2) - R_k^{(2)} *_2 (bf)(x_1, x_2). \tag{1-3}$$

Iterated commutators arise in the study of commutators of multiparameter singular integral operators which are tensor products. In the flag setting, our iterated commutator takes the following form:

Definition 1.1. Given two functions $b, f \in L^2(\mathbb{R}^{n+m})$, the iterated commutator in the flag setting of $\mathbb{R}^n \times \mathbb{R}^m$ is

$$\begin{aligned}[][b, R_j^{(1)}], R_k^{(2)}]_2(f) := b(x_1, x_2) R_j^{(1)} * R_k^{(2)} *_2 f(x_1, x_2) - R_j^{(1)} * (b \cdot R_k^{(2)} *_2 f)(x_1, x_2) \\ - R_k^{(2)} *_2 (b \cdot R_j^{(1)} * f)(x_1, x_2) + R_k^{(2)} *_2 R_j^{(1)} * (b \cdot f)(x_1, x_2).\end{aligned}$$

We point out that another possible definition via $[[b, R_k^{(2)}]_2, R_j^{(1)}](f)$ turns out to be equivalent; see Proposition 2.5 in Section 2.

We also introduce the *big* commutator in the flag setting as follows.

Definition 1.2. Given two functions $b, f \in L^2(\mathbb{R}^{n+m})$, the big commutator in the flag setting of $\mathbb{R}^n \times \mathbb{R}^m$ is

$$[b, R_{j,k}](f)(x) := b(x) R_{j,k}(f)(x) - R_{j,k}(bf)(x). \tag{1-4}$$

The main results, below, of this paper relate iterated and big commutator bounds to flag BMO spaces. As the definition of the space $\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ is very technical, we refer the reader to Section 2, Definition 2.4 for details.

Theorem 1.3. *Suppose $b \in \mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ and $1 < p < \infty$. Then for every $j = 1, \dots, n+m$, $k = 1, \dots, m$, $f \in L^p(\mathbb{R}^{n+m})$,*

$$\|[[b, R_j^{(1)}], R_k^{(2)}]_2(f)\|_{L^p(\mathbb{R}^{n+m})} \lesssim \|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)} \|f\|_{L^p(\mathbb{R}^{n+m})}. \tag{1-5}$$

Lacking methods related to analyticity ([Ferguson and Sadosky 2000] for the Hilbert transform) or wavelets [Lacey et al. 2009; 2010; Dalenc and Ou 2016], we instead obtain this upper bound using the duality argument and the tools of semigroups, harmonic function extensions and techniques from multiparameter analysis.

Next, we introduce the little flag BMO space. The flag structure has a geometry which is reflected by the axes-parallel rectangles $R = I \times J \subset \mathbb{R}^{n+m}$ satisfying $\ell(I) \le \ell(J)$, the collection of which is referred to as *flag rectangles*, denoted by $\mathcal{R}_{\mathcal{F}}$. One can then define the little flag BMO space and the flag-type Muckenhoupt weights $A_{\mathcal{F},p}$ with respect to $\mathcal{R}_{\mathcal{F}}$.

Definition 1.4. A locally integrable function b is in *little flag BMO space*, denoted by $\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$, if

$$\|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)} := \sup_{R\in\mathcal{R}_{\mathcal{F}}} \frac{1}{|R|} \int_R |b(x, y) - \langle b\rangle_R| \, dx \, dy < \infty, \tag{1-6}$$

where $\langle b\rangle_R = (1/|R|) \int_R b(x_1, x_2) \, dx_1 \, dx_2$.

Theorem 1.5. *Suppose $T_{\mathcal{F}}$ is a flag singular integral operator on $\mathbb{R}^n \times \mathbb{R}^m$, $b \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ and $1 < p < \infty$. Then for $f \in L^p(\mathbb{R}^{n+m})$,*

$$\|[b, T_{\mathcal{F}}](f)\|_{L^p(\mathbb{R}^{n+m})} \lesssim \|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)} \|f\|_{L^p(\mathbb{R}^{n+m})}. \tag{1-7}$$

In the above, the flag singular integral $T_{\mathcal{F}}$ can be taken as the Riesz transform $R_{j,k}$. The class of flag singular integral operators $T_{\mathcal{F}}$ naturally generalize the Riesz transforms $R_{j,k}$ and are assumed to be associated to kernels having a standard flag structure. We refer the reader to Definition 4.4 in Section 4 for its precise definition. To obtain this upper bound, we study the little flag BMO space $\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ and find the connection with the John–Nirenberg BMO space on $\mathbb{R}^{n+m}$ and on $\mathbb{R}^m$. We also establish the bridge between functions in $\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ and weights in $A_{\mathcal{F},p}$. These structures lead to the upper bound for $[b, R_{j,k}](f)$.

As application, the commutator estimates obtained above imply certain versions of div-curl lemmas, which seem to be first of their kind in the flag setting. Roughly speaking, a div-curl lemma says that if vector fields E and B initially in L^2 have some cancellation (e.g., divergence or curl zero) then one can expect their dot product $E \cdot B$ to belong to a better space of functions instead of just L^1 (as provided for by Cauchy–Schwarz). The cancellation conditions allow one to deduce some type of cancellation, e.g., $\int E \cdot B = 0$, suggesting that the function should belong to a suitable Hardy space since it is integrable and has mean zero. The algebraic structure of $E \cdot B$ coupled with the duality between Hardy spaces and BMO spaces then points to the use of the commutator theorem to arrive at the membership of $E \cdot B$ in the Hardy space; different commutator results suggest different div-curl lemmas that can be explored. In the classical one-parameter setting, the div-curl lemma says that given two vector fields, one with divergence zero and the other with curl zero, their dot product belongs to a Hardy space [Coifman et al. 1993]. Later on, Lacey, Petermichl, and the fourth and the fifth authors proved multiple versions of div-curl lemmas in the multiparameter setting [Lacey et al. 2012], which are expected since the multiparameter setting offers several different interpretations of the Hardy and BMO spaces. Thus, it is natural that our Theorems 1.3 and 1.5 lead to two versions of flag-type div-curl lemmas.

First, consider vector fields on $\mathbb{R}^n \times \mathbb{R}^m$ that take values in $\mathcal{M}_{n+m,m}$ and are associated with the flag structure (see Section 5 for the precise definitions and details). We establish the div-curl lemma in the flag setting with respect to the flag Hardy space below, which is a consequence of Theorem 1.3.

Theorem 1.6. *Let $1 < p, q < \infty$ with $1/p + 1/q = 1$. Suppose that E, B are vector fields on $\mathbb{R}^n \times \mathbb{R}^m$ taking the values in $\mathcal{M}_{n+m,m}$, associated with the flag structure. Moreover, suppose $E = E^{(1)} *_2 E^{(2)} \in L^p_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m; \mathcal{M}_{n+m,m})$ and $B = B^{(1)} *_2 B^{(2)} \in L^q_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m; \mathcal{M}_{n+m,m})$ satisfy*

$$\operatorname{div}_{(x,y)} E_j^{(1)}(x, y) = 0 \quad \textit{and} \quad \operatorname{curl}_{(x,y)} B_j^{(1)}(x, y) = 0 \quad \textit{for all } k$$

and

$$\operatorname{div}_y E_k^{(2)}(x, y) = 0 \quad \textit{and} \quad \operatorname{curl}_y B_k^{(2)}(x, y) = 0 \quad \textit{for all } x \in \mathbb{R}^n, \textit{ for all } j.$$

Then $E \cdot B$ belongs to the flag Hardy space $H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ with

$$\|E \cdot B\|_{H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim \|E\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m; \mathcal{M}_{n+m,m})} \|B\|_{L^q(\mathbb{R}^n \times \mathbb{R}^m; \mathcal{M}_{n+m,m})}. \tag{1-8}$$

We also prove another version of the div-curl lemma in the flag setting, which is with respect to the Hardy spaces on $\mathbb{R}^{n+m}$ and on $\mathbb{R}^m$, respectively. This version relies on the intermediate result in the proof of Theorem 1.5, namely, the structure of the flag little bmo space.

Theorem 1.7. *Let $1 < p, q < \infty$ with $1/p + 1/q = 1$. Suppose that E, B are vector fields on $\mathbb{R}^n \times \mathbb{R}^m$ taking the values in $\mathbb{R}^{n+m}$. Moreover, suppose $E \in L^p(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^{n+m})$ and $B \in L^q(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^{n+m})$ satisfy*

$$\operatorname{div}_{(x,y)} E(x, y) = 0 \quad \textit{and} \quad \operatorname{curl}_{(x,y)} B(x, y) = 0$$

and

$$\operatorname{div}_y E(x, y) = 0 \quad \textit{and} \quad \operatorname{curl}_y B(x, y) = 0 \quad \textit{for all } x \in \mathbb{R}^n.$$

Then we have

$$\|E \cdot B\|_{H^1(\mathbb{R}^{n+m})} \lesssim \|E\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^{n+m})} \|B\|_{L^q(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^{n+m})}, \tag{1-9}$$

and

$$\int_{\mathbb{R}^m} \|E(\,\cdot\,, y) \cdot_2 B(\,\cdot\,, y)\|_{H^1(\mathbb{R}^m)}\, dy \lesssim \|E\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^{n+m})} \|B\|_{L^q(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^{n+m})}, \tag{1-10}$$

where

$$E(x, y) \cdot_2 B(x, y) := \sum_{k=1}^{m} E_{n+k}(x, y) B_k(x, y).$$

It is known that the div-curl lemma in the classical setting has many applications in PDE and compensated compactness [Coifman et al. 1993]. Similarly, we expect that the flag-type div-curl lemmas described above would have interesting implications in these directions as well. For instance, following the ideas in [Coifman et al. 1993], one can study weak convergence problems in the flag Hardy space. And it would be interesting to know whether one can use the flag-type regularity (implied by our div-curl lemmas) of certain nonlinear quantities to obtain improved regularity results for certain nonlinear PDE.

This paper is organized as follows. In Section 2 we provide necessary preliminaries with respect to the flag structures. In Section 3 we study the flag iterated commutators as in Definition 1.1 and prove Theorem 1.3. In Section 4 we give a complete treatment of the flag little bmo spaces and flag-type Muckenhoupt A_p weights, toward the proof of Theorem 1.5. In the last section, we apply the boundedness of flag commutators from Theorems 1.3 and 1.5 to establish the flag div-curl results, Theorems 1.6 and 1.7.

2. Preliminaries in the flag setting

Recall the classical Poisson kernel on $\mathbb{R}^n$:

$$P(x) := \frac{c_n}{(1+|x|^2)^{(n+1)/2}}.$$

And we define

$$P_t(x) := \frac{1}{t^n} P\left(\frac{x}{t}\right).$$

For $f \in L^1(\mathbb{R}^n)$, let $F(x,t) := P_t * f(x)$. Then we have the following standard pointwise estimates for the Poisson integral; see in particular [Stein 1993].

Proposition 2.1. *Suppose $f \in L^1(\mathbb{R}^n)$. Then*

$$\sup_{(x,t)\in\mathbb{R}^{n+1}_+} t^{n+k} |\nabla^k_{x,t} F(x,t)| \le C \|f\|_{L^1(\mathbb{R}^n)}. \tag{2-1}$$

We now recall the flag Poisson kernel given by

$$P(x,y) = P^{(1)} *_{\mathbb{R}^m} P^{(2)}(x,y) = \int_{\mathbb{R}^m} P^{(1)}(x, y-z) P^{(2)}(z)\, dz,$$

where

$$P^{(1)}(x,y) = \frac{c_{n+m}}{(1+|x|^2+|y|^2)^{(n+m+1)/2}} \quad \text{and} \quad P^{(2)}(z) = \frac{c_m}{(1+|z|^2)^{(m+1)/2}}$$

are the classical Poisson kernels on $\mathbb{R}^{n+m}$ and $\mathbb{R}^m$, respectively. Then we have

$$P_{t_1,t_2}(x,y) = P^{(1)}_{t_1} *_{\mathbb{R}^m} P^{(2)}_{t_2}(x,y).$$

We define the Lusin area function with respect to $u = P_{t_1,t_2} * f$ as follows.

Definition 2.2. For $f \in L^1(\mathbb{R}^n \times \mathbb{R}^m)$ and $u(x_1,x_2,t_1,t_2) = P_{t_1,t_2} * f(x_1,x_2)$, the Lusin area integral of $u(x_1,x_2,t_1,t_2)$, denoted by $S_{\mathcal{F}}(u)$, is defined by

$$S_{\mathcal{F}}(u)(x_1,x_2) = \left\{ \int_{\mathbb{R}^{n+1}_+} \int_{\mathbb{R}^{m+1}_+} \chi_{t,s}(x_1-w_1, x_2-w_2) |t_1 \nabla^{(1)} t_2 \nabla^{(2)} u(w_1,w_2,t_1,t_2)|^2 \frac{dw_1\,dt_1}{t_1^{n+m+1}} \frac{dw_2\,dt_2}{t_2^{m+1}} \right\}^{\frac{1}{2}},$$

where $\nabla^{(1)} = (\partial_{t_1}, \partial_{w_{1,1}} \cdots \partial_{w_{1,n}}, \partial_{w_{2,1}} \cdots \partial_{w_{2,m}})$ is the standard gradient on $\mathbb{R}^{n+m+1}$, and $\nabla^{(2)} = (\partial_{t_2}, \partial_{w_{2,1}} \cdots \partial_{w_{2,m}})$ is the standard gradient on $\mathbb{R}^{m+1}$, and

$$\chi_{t_1,t_2}(x_1, x_2) := \chi^{(1)}_{t_1} *_{\mathbb{R}^m} \chi^{(2)}_{t_2}(x_1, x_2), \tag{2-2}$$

$\chi^{(1)}_{t_1}(x_1, x_2) := t_1{}^{-(n+m)} \chi^{(1)}(x_1/t_1, x_2/t_1)$, $\chi^{(2)}_{t_2}(z) := t_2{}^{-m} \chi^{(2)}(z/t_2)$, and $\chi^{(1)}(x, y)$ and $\chi^{(2)}(z)$ are the indicator functions of the unit balls of $\mathbb{R}^{n+m}$ and $\mathbb{R}^m$, respectively.

Definition 2.3. The flag Hardy space $H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ is defined to be the collection of $f \in L^1(\mathbb{R}^n \times \mathbb{R}^m)$ such that $S_{\mathcal{F}}(u) \in L^1(\mathbb{R}^n \times \mathbb{R}^m)$. The norm of $H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ is defined by

$$\|f\|_{H^1_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)} = \|S_{\mathcal{F}}(u)\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)}. \tag{2-3}$$

We now recall the definition of the flag BMO space.

Definition 2.4. The flag BMO space $\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ is defined to be the collection of $b \in L^1_{\mathrm{loc}}(\mathbb{R}^n \times \mathbb{R}^m)$ such that

$$\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)} := \sup_{\Omega} \left(\frac{1}{|\Omega|} \int_{T(\Omega)} |t_1 \nabla^{(1)} t_2 \nabla^{(2)} u(w_1, w_2, t_1, t_2)|^2 \, \frac{dw_1\, dt_1 dw_2\, dt_2}{t_1 t_2} \right)^{\frac{1}{2}} < \infty, \tag{2-4}$$

where the supremum is taken over all open sets in $\mathbb{R}^n \times \mathbb{R}^m$ with finite measures, and $T(\Omega) = \bigcup_{R \subset \Omega} T(R)$ with $R = I \times J$, $\ell(I) \le \ell(I)$ and $T(R) = I \times \left(\frac{1}{2}\ell(I), \ell(I)\right] \times J \times \left(\frac{1}{2}\ell(J), \ell(J)\right]$.

Proposition 2.5. *Given two functions b, $f \in L^2(\mathbb{R}^{n+m})$, we have*

$$[[b, R^{(1)}_j], R^{(2)}_k]_2(f) = [[b, R^{(2)}_k]_2, R^{(1)}_j](f). \tag{2-5}$$

Proof. By definition, we see that

$$\begin{aligned} [[b, R^{(1)}_j], R^{(2)}_k]_2(f)(x_1, x_2) &= [b, R^{(1)}_j] R^{(2)}_k *_2 f(x_1, x_2) - R^{(2)}_k *_2 ([b, R^{(1)}_j](f))(x_1, x_2) \\ &= b(x_1, x_2) R^{(1)}_j * R^{(2)}_k *_2 f(x_1, x_2) - R^{(1)}_j * (b \cdot R^{(2)}_k *_2 f)(x_1, x_2) \\ &\qquad - R^{(2)}_k *_2 (b \cdot R^{(1)}_j * f - R^{(1)}_j * (b \cdot f))(x_1, x_2) \\ &= b(x_1, x_2) R^{(1)}_j * R^{(2)}_k *_2 f(x_1, x_2) - R^{(1)}_j * (b \cdot R^{(2)}_k *_2 f)(x_1, x_2) \\ &\qquad - R^{(2)}_k *_2 (b \cdot R^{(1)}_j * f)(x_1, x_2) + R^{(2)}_k *_2 R^{(1)}_j * (b \cdot f)(x_1, x_2). \end{aligned}$$

And we also have

$$\begin{aligned} [[b, R^{(2)}_k]_2, R^{(1)}_j](f)(x_1, x_2) &= [b, R^{(2)}_k]_2 R^{(1)}_j * f(x_1, x_2) - R^{(1)}_j * ([b, R^{(2)}_k]_2(f))(x_1, x_2) \\ &= b(x_1, x_2) R^{(2)}_k *_2 R_j * f(x_1, x_2) - R^{(2)}_k *_2 (b \cdot R^{(1)}_j * f)(x_1, x_2) \\ &\qquad - R^{(1)}_j * (b \cdot R^{(2)}_k *_2 f - R^{(2)}_k *_2 (b \cdot f))(x_1, x_2) \\ &= b(x_1, x_2) R^{(2)}_k *_2 R^{(1)}_j * f(x_1, x_2) - R^{(2)}_k *_2 (b \cdot R^{(1)}_j * f)(x_1, x_2) \\ &\qquad - R^{(1)}_j * (b \cdot R^{(2)}_k *_2 f)(x_1, x_2) + R^{(1)}_j * R^{(2)}_k *_2 (b \cdot f)(x_1, x_2). \end{aligned}$$

It is direct to see that, by changing of variables,

$$\begin{aligned} R_k^{(2)} *_2 R_j^{(1)} * f(x_1, x_2) &= \int R_k^{(2)}(x_2 - z) R_j^{(1)}(x_1 - y_1, z - y_2) f(y_1, y_2)\, dz\, dy_1\, dy_2 \\ &= \int R_k^{(2)}(\tilde{z} - y_2) R_j^{(1)}(x_1 - y_1, x_2 - \tilde{z}) f(y_1, y_2)\, d\tilde{z}\, dy_1\, dy_2 \\ &= \int R_j^{(1)}(x_1 - y_1, x_2 - \tilde{z}) R_k^{(2)}(\tilde{z} - y_2) f(y_1, y_2)\, d\tilde{z}\, dy_1\, dy_2 \\ &= R_j^{(1)} * R_k^{(2)} *_2 f(x_1, x_2), \end{aligned}$$

which implies that (2-5) holds. □

3. Upper bound of the iterated commutator $[[b, R_i^{(1)}], R_j^{(2)}]_2$

In this section, we prove Theorem 1.3, i.e., the upper bound of the iterated commutator $[[b, R_i^{(1)}], R_j^{(2)}]_2$. As we pointed out earlier, in the flag setting, there is lack of a suitable wavelet basis or Haar basis and hence the approaches in [Lacey et al. 2009; Dalenc and Ou 2016] are not available. We establish a fundamental duality argument (Lemma 3.3) with respect to general flag-type area integrals and flag Carleson measures, and then apply the technique of harmonic expansion to obtain the full versions of flag-type Carleson measure inequalities (Proposition 3.5), which plays the role of "paraproducts". Then, by considering the bilinear form associated with the iterated commutator $[[b, R_i^{(1)}], R_j^{(2)}]_2$ and by integration by parts, we can decompose the bilinear form into a summation of different versions of "paraproducts". Then the upper bound of the iterated commutator $[[b, R_i^{(1)}], R_j^{(2)}]_2$ follows from applying Proposition 3.5 to each "paraproducts".

Extension via flag Poisson operator. For any $f \in L^1(\mathbb{R}^n \times \mathbb{R}^m)$, we define the flag Poisson integral of f by

$$F(x_1, x_2, t_1, t_2) := P_{t_1,t_2} * f(x_1, y_2), \tag{3-1}$$

where

$$P_{t_1,t_2}(x_1, x_2) = P_{t_1}^{(1)} *_{\mathbb{R}^m} P_{t_2}^{(2)}(x_1, x_2). \tag{3-2}$$

Since $P(x_1, x_2) \in L^1(\mathbb{R}^n \times \mathbb{R}^m)$, it easy to see that $F(x_1, x_2, t_1, t_2)$ is well-defined. Moreover, for any fixed t_1 and t_2, we know $P_{t_1,t_2} * f(x_1, x_2)$ is a bounded C^∞ function and the function $F(x_1, x_2, t_1, t_2)$ is harmonic in (x_1, x_2, t_1) and (x_2, t_2), respectively. $F(x_1, x_2, t_1, t_2)$ is the flag harmonic extension of f to $\mathbb{R}_+^{n+1} \times \mathbb{R}_+^{m+1}$. More precisely,

$$\begin{aligned} \Delta_{\mathbb{R}^{n+m+1}} F(x_1, x_2, t_1, t_2) &= (\partial_{t_1}^2 + \Delta_{x_1,x_2}) F(x_1, x_2, t_1, t_2) = 0 \quad \text{in } \mathbb{R}_+^{n+m+1}, \\ \Delta_{\mathbb{R}^{m+1}} F(x_1, x_2, t_1, t_2) &= (\partial_{t_2}^2 + \Delta_{x_2}) F(x_1, x_2, t_1, t_2) = 0 \quad \text{in } \mathbb{R}_+^{m+1}, \end{aligned} \tag{3-3}$$

and

$$\begin{aligned} \lim_{t_1 \to 0} \partial_{t_1} F(x_1, x_2, t_1, t_2) &= -(\Delta_{x_1,x_2})^{\frac{1}{2}} P^{(2)} *_{\mathbb{R}^m} f(x_1, x_2) \quad \text{on } \mathbb{R}^{n+m}, \\ \lim_{t_2 \to 0} \partial_{t_2} F(x_1, x_2, t_1, t_2) &= -(\Delta_{x_2})^{\frac{1}{2}} P^{(1)} * f(x_1, x_2) \quad \text{on } \mathbb{R}^{n+m}, \end{aligned}$$

$$\begin{aligned}
&\lim_{t_1\to 0} F(x_1,x_2,t_1,t_2) = P^{(2)} *_{\mathbb{R}^m} f(x_1,x_2) && \text{on } \mathbb{R}^{n+m},\\
&\lim_{t_2\to 0} F(x_1,x_2,t_1,t_2) = P^{(1)} * f(x_1,x_2) && \text{on } \mathbb{R}^{n+m},\\
&\lim_{t_1\to 0,\ t_2\to 0} F(x_1,x_2,t_1,t_2) = f(x_1,x_2) && \text{on } \mathbb{R}^{n+m},\\
&\lim_{|(x_1,x_2,t_1)|\to\infty} F(x_1,x_2,t_1,t_2) = 0,\\
&\lim_{|(x_2,t_2)|\to\infty} F(x_1,x_2,t_1,t_2) = 0.
\end{aligned}$$

We then have the following lemma providing a connection between the boundary values f and the flag harmonic extension F. This follows from the decay of the flag harmonic extensions of f and repeated applications of integration by parts in the variables t_1 and t_2.

Lemma 3.1. *For $f \in L^1(\mathbb{R}^n \times \mathbb{R}^m)$, let F be the same as in* (3-1)*. Then we have*

$$\int_{\mathbb{R}^n\times\mathbb{R}^m} f(x_1,x_2)\,dx_1\,dx_2 = \int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 F(x_1,x_2,t_1,t_2)\,dx_1\,dx_2\,dt_1\,dt_2. \tag{3-4}$$

Proof. We start from the right-hand side of (3-4). We write

$$\begin{aligned}
\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 F(x_1,x_2,t_1,t_2)\,dx_1\,dx_2\,dt_1\,dt_2\\
&= \int_{\mathbb{R}^{m+1}_+} t_2\,\partial_{t_2}^2 P^{(2)}_{t_2} *_{\mathbb{R}^m} \biggl(\int_{\mathbb{R}^{n+1}_+} t_1\,\partial_{t_1}^2 P^{(1)}_{t_1} * f(x_1,x_2)\,dx_1\,dt_1\biggr)\,dx_2\,dt_2\\
&= \int_{\mathbb{R}^m} \biggl(\int_{\mathbb{R}^{n+1}_+} t_1\,\partial_{t_1}^2 P^{(1)}_{t_1} * f(x_1,x_2)\,dx_1\,dt_1\biggr)\,dx_2,
\end{aligned}$$

where the last equality follows from decay of the flag harmonic extensions of f and using integration by parts in the variable t_2. To continue, we write the right-hand side of the last equality above as

$$\int_{\mathbb{R}^{n+m+1}_+} t_1\,\partial_{t_1}^2 P^{(1)}_{t_1} * f(x_1,x_2)\,dx_1\,dx_2\,dt_1 = \int_{\mathbb{R}^{n+m}} f(x_1,x_2)\,dx_1\,dx_2,$$

which yields (3-4). Again, the last equality follows from decay of the flag harmonic extensions of f and using integration by parts in the variable t_1. □

Flag area functions and estimates. We also have a more general version of the area function.

Definition 3.2. For a function $G(x_1,x_2,t_1,t_2)$ defined on $\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+$, the general flag-type Lusin area integral of G is defined by

$$S_{\mathcal{F},L}(G)(x_1,x_2) := \biggl\{\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+} \chi_{t,s}(x_1-w_1,x_2-w_2)|G(w_1,w_2,t_1,t_2)|^2\,\frac{dw_1\,dt_1}{t_1^{n+m+1}}\,\frac{dw_2\,dt_2}{t_2^{m+1}}\biggr\}^{\frac12}. \tag{3-5}$$

Lemma 3.3. *Suppose $F(x_1, x_2, t_1, t_2)$ and $G(x_1, x_2, t_1, t_2)$ are defined on $\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+$. Then the following estimate holds*:

$$\begin{aligned}
&\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+} F(x_1, x_2, t_1, t_2) G(x_1, x_2, t_1, t_2)\; dx_1\, dx_2\, dt_1\, dt_2 \\
&\quad \le C \sup_{\Omega\subset\mathbb{R}^n\times\mathbb{R}^m} \biggl(\frac{1}{|\Omega|}\int_{T(\Omega)} t_1\, t_2\, |F(y_1, y_2, t_1, t_2)|^2\, dy_1\, dy_2\, dt_1\, dt_2\biggr)^{1/2} \\
&\quad\quad \times \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \biggl(\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+} \chi_{t_1,t_2}(x_1 - y_1, x_2 - y_2) |G(y_1, y_2, t_1, t_2)|^2 \frac{dy_1\, dy_2\, dt_1\, dt_2}{t_1^{n+m+1} t_2^{m+1}}\biggr)^{1/2} dx_1\, dx_2. \quad (3\text{-}6)
\end{aligned}$$

Proof. Suppose both factors on the right-hand side above are finite, since otherwise there is nothing to prove. We also note that the second factor is actually $\|S_{\mathcal{F}}(G)\|_{L^1(\mathbb{R}^n\times\mathbb{R}^m)}$.

We now let

$$\Omega_k := \{(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^m :\ S_{\mathcal{F},L}(G)(x_1, x_2) > 2^k\}$$

and define

$$B_k := \bigl\{R = I_1 \times I_2 :\ |(I_1 \times I_2) \cap \Omega_k| > \tfrac{1}{2}|I_1 \times I_2|,\ |(I_1 \times I_2) \cap \Omega_{k+1}| \le \tfrac{1}{2}|I_1 \times I_2|\bigr\},$$

where I_1 and I_2 are dyadic cubes in $\mathbb{R}^n$ and $\mathbb{R}^m$ with side-lengths $\ell(I)$ and $\ell(J)$ satisfying $\ell(I) \le \ell(J)$. Moreover, we define

$$\Omega_k = \bigcup_{R\in B_k} R \quad\text{and}\quad \widetilde{\Omega}_k = \bigl\{(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^m :\ M_{\text{flag}}(\chi_{\Omega_k})(x_1, x_2) > \tfrac{1}{2}\bigr\}.$$

Next, we have

$$\begin{aligned}
&\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+} F(x_1, x_2, t_1, t_2) G(x_1, x_2, t_1, t_2)\; dx_1\, dx_2\, dt_1\, dt_2 \\
&\quad = \sum_k \sum_{R\in B_k} \int_{T(R)} \sqrt{t_1t_2} F(x_1, x_2, t_1, t_2)\, \frac{G(x_1, x_2, t_1, t_2)}{\sqrt{t_1t_2}}\, dx_1\, dx_2\, dt_1\, dt_2 \\
&\quad \le \sum_k \biggl(\sum_{R\in B_k} \int_{T(R)} t_1t_2 |F(x_1, x_2, t_1, t_2)|^2\, dx_1\, dx_2\, dt_1\, dt_2\biggr)^{1/2} \\
&\qquad\qquad \times \biggl(\sum_{R\in B_k} \int_{T(R)} |G(x_1, x_2, t_1, t_2)|^2\, \frac{dx_1\, dx_2\, dt_1\, dt_2}{t_1t_2}\biggr)^{1/2} \\
&\quad = \sum_k \biggl(\frac{1}{|\Omega_k|} \sum_{R\in B_k} \int_{T(R)} t_1t_2 |F(x_1, x_2, t_1, t_2)|^2\, dx_1\, dx_2\, dt_1\, dt_2\biggr)^{1/2} \\
&\qquad\qquad \times \biggl(|\Omega_k| \sum_{R\in B_k} \int_{T(R)} |G(x_1, x_2, t_1, t_2)|^2\, \frac{dx_1\, dx_2\, dt_1\, dt_2}{t_1t_2}\biggr)^{1/2} \\
&\quad \le \sum_k \biggl(\frac{1}{|\Omega_k|} \int_{T(\Omega_k)} t_1t_2 |F(x_1, x_2, t_1, t_2)|^2\, dx_1\, dx_2\, dt_1\, dt_2\biggr)^{1/2} \\
&\qquad\qquad \times \biggl(|\widetilde{\Omega}_k| \sum_{R\in B_k} \int_{T(R)} |G(x_1, x_2, t_1, t_2)|^2\, \frac{dx_1\, dx_2\, dt_1\, dt_2}{t_1t_2}\biggr)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&\le \sup_{\Omega\subset\mathbb{R}^n\times\mathbb{R}^m}\left(\frac{1}{|\Omega|}\int_{T(\Omega)} t_1t_2|F(x_1,x_2,t_1,t_2)|^2\,dx_1\,dx_2\,dt_1\,dt_2\right)^{1/2}\\
&\qquad\qquad\times\sum_k\left(|\widetilde{\Omega}_k|\sum_{R\in B_k}\int_{T(R)}|G(x_1,x_2,t_1,t_2)|^2\,\frac{dx_1\,dx_2\,dt_1\,dt_2}{t_1t_2}\right)^{1/2}.
\end{aligned}
$$

As for the second factor in the last inequality above, note that

$$
\begin{aligned}
2^{2k}|\widetilde{\Omega}_k\backslash\Omega_k| &\ge \int_{\widetilde{\Omega}_k\backslash\Omega_k} S_{\mathcal{F},L}(G)(x_1,x_2)^2\,dx_1\,dx_2\\
&=\int_{\widetilde{\Omega}_k\backslash\Omega_k}\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+}\chi_{t_1,t_2}(x_1-y_1,x_2-y_2)|G(y_1,y_2,t_1,t_2)|^2\,\frac{dy_1\,dy_2\,dt_1\,dt_2}{t_1^{n+m+1}t_2^{m+1}}\,dx_1\,dx_2\\
&=\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+}\int_{\widetilde{\Omega}_k\backslash\Omega_k}\chi_{t_1,t_2}(x_1-y_1,x_2-y_2)\,dx_1\,dx_2\,|G(y_1,y_2,t_1,t_2)|^2\,\frac{dy_1\,dy_2\,dt_1\,dt_2}{t_1^{n+m+1}t_2^{m+1}}\\
&\approx\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+}|G(y_1,y_2,t_1,t_2)|^2\,\frac{dy_1\,dy_2\,dt_1\,dt_2}{t_1t_2}\\
&\ge\sum_{R\in B_k}\int_{T(R)}|G(x_1,x_2,t_1,t_2)|^2\,\frac{dx_1\,dx_2\,dt_1\,dt_2}{t_1t_2}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+}F(x_1,x_2,t_1,t_2)G(x_1,x_2,t_1,t_2)\,dx_1\,dx_2\,dt_1\,dt_2\\
&\qquad\le\sup_{\Omega\subset\mathbb{R}^n\times\mathbb{R}^m}\left(\frac{1}{|\Omega|}\int_{T(\Omega)}t_1t_2|F(x_1,x_2,t_1,t_2)|^2\,dx_1\,dx_2\,dt_1\,dt_2\right)^{1/2}\sum_k(|\widetilde{\Omega}_k|2^{2k}|\widetilde{\Omega}_k\backslash\Omega_k|)^{1/2}\\
&\qquad\le\sup_{\Omega\subset\mathbb{R}^n\times\mathbb{R}^m}\left(\frac{1}{|\Omega|}\int_{T(\Omega)}|t_1t_2F(x_1,x_2,t_1,t_2)|^2\,\frac{dx_1\,dx_2\,dt_1\,dt_2}{t_1t_2}\right)^{1/2}\sum_k|\Omega_k|2^k\\
&\qquad\le\sup_{\Omega\subset\mathbb{R}^n\times\mathbb{R}^m}\left(\frac{1}{|\Omega|}\int_{T(\Omega)}|t_1t_2F(x_1,x_2,t_1,t_2)|^2\,\frac{dx_1\,dx_2\,dt_1\,dt_2}{t_1t_2}\right)^{1/2}\|S_{\mathcal{F},L}(G)\|_{L^1(\mathbb{R}^n\times\mathbb{R}^m)},
\end{aligned}
$$

which gives (3-6). ☐

From Lemma 3.3 above and the definition of $\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$, we can obtain the following corollary immediately.

Corollary 3.4. *Suppose $G(x_1,x_2,t_1,t_2)$ is defined on $\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+$, and $F(x_1,x_2,t_1,t_2):=P_{t_1,t_2}*f(x_1,x_2)$, where $f\in\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$. Then we have*

$$
\begin{aligned}
&\int_{\mathbb{R}^{n+1}_+}\int_{\mathbb{R}^{m+1}_+}|\nabla^{(1)}\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,|G(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2\\
&\qquad\qquad\le C\|f\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|S_{\mathcal{F},L}(G)\|_{L^1(\mathbb{R}^n\times\mathbb{R}^m)}. \qquad (3\text{-}7)
\end{aligned}
$$

Moreover, based on Lemma 3.3, we can also establish the following estimates.

Proposition 3.5. *Suppose* $F(x_1,x_2,t_1,t_2)=P_{t_1,t_2}*f(x_1,x_2)$, $G(x_1,x_2,t_1,t_2)=P_{t_1,t_2}*g(x_1,x_2)$, *and* $B(x_1,x_2,t_1,t_2)=P_{t_1,t_2}*b(x_1,x_2)$. *Then we have*

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(1)}\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-8}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(1)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-9}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-10}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-11}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(1)}\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-12}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(1)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-13}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\leq C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-14}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla_{x_1,x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-15}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |\nabla^{(1)}\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-16}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |\nabla^{(1)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-17}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-18}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-19}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |\nabla^{(1)}\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-20}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|\,|\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times |\nabla^{(1)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2 \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-21}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-22}$$

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}. \tag{3-23}$$

Proof. We first point out that for $f\in C_0^\infty(\mathbb{R}^{n+m})$, $F(x_1,x_2,t_1,t_2)=P_{t_1,t_2}*f(x_1,x_2)$,

$$\sup_{\substack{(y_1,y_2,t_1,t_2)\\ \chi_{t_1,t_2}(x_1-y_1,x_2-y_2)\ne0}}|F(y_1,y_2,t_1,t_2)|\le \sup_{\substack{(y_1,y_2,t_1,t_2)\\ |x_1-y_1|<t_1+t_2,|x_2-y_2|<t_2}}|P_{t_1,t_2}*f(y_1,y_2)|$$
$$\le M_1(M_2(f(\cdot_1,\cdot))(\cdot_2))(x_1,x_2),$$

where M_1 and M_2 are the Hardy–Littlewood maximal functions on $\mathbb{R}^{n+m}$ and $\mathbb{R}^m$, respectively.

Next, based on the estimate above and from the property of the Poisson semigroup, we have

$$\sup_{\substack{(y_1,y_2,t_1,t_2)\\ \chi_{t_1,t_2}(x_1-y_1,x_2-y_2)\ne0}}|\partial_{t_1}\partial_{t_2}F(y_1,y_2,t_1,t_2)|\le \sup_{\substack{(y_1,y_2,t_1,t_2)\\ |x_1-y_1|<t_1+t_2,|x_2-y_2|<t_2}}|P_{t_1,t_2}*((-\Delta_{(1)})^{\frac12}(-\Delta_{(2)})^{\frac12}f)(y_1,y_2)|$$
$$\le M_1\big(M_2(((-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f)(\cdot_1,\cdot))(\cdot_2)\big)(x_1,x_2).$$

Also, we have

$$\sup_{\substack{(y_1,y_2,t_1,t_2)\\ \chi_{t_1,t_2}(x_1-y_1,x_2-y_2)\ne0}}|\nabla_{y_1,y_2}\nabla_{y_2}F(y_1,y_2,t_1,t_2)|\le \sup_{\substack{(y_1,y_2,t_1,t_2)\\ |x_1-y_1|<t_1+t_2,|x_2-y_2|<t_2}}|P_{t_1,t_2}*(\nabla_{\cdot_1,\cdot_2}\nabla_{\cdot_2}f)(y_1,y_2)|$$
$$\le M_1\big(M_2((\nabla_{\cdot_1,\cdot_2}\nabla_{\cdot_2}f)(\cdot_1,\cdot))(\cdot_2)\big)(x_1,x_2).$$

Then, we first consider (3-8). Based on the estimates above and Corollary 3.4, we have

$$\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} t_1t_2|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)||\nabla_{x_1,x_2}\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G(x_1,x_2,t_1,t_2)| \times|\nabla^{(1)}\nabla^{(2)}F(x_1,x_2,t_1,t_2)|\,dx_1\,dx_2\,dt_1\,dt_2$$
$$\le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\int_{\mathbb{R}^n\times\mathbb{R}^m} S_{\mathcal{F},L}(t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}\nabla^{(1)}\nabla^{(2)}G)(x_1,x_2) \times\Big(M_1\big(M_2(((-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f)(\cdot_1,\cdot))(\cdot_2)\big)(x_1,x_2) +M_1\big(M_2((\nabla_{\cdot_1,\cdot_2}\nabla_{\cdot_2}f)(\cdot_1,\cdot))(\cdot_2)\big)(x_1,x_2)\Big)\,dx_1\,dx_2$$
$$\le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)} \int_{\mathbb{R}^n\times\mathbb{R}^m} S_{\mathcal{F}}\big(\nabla_{x_1,x_2}\nabla_{x_2}(-\Delta_{x_1,x_2})^{-\frac12}(-\Delta_{x_2})^{-\frac12}(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}G\big)(x_1,x_2) \times\Big(M_1\big(M_2(((-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f)(\cdot_1,\cdot))(\cdot_2)\big)(x_1,x_2) +M_1\big(M_2((\nabla_{\cdot_1,\cdot_2}\nabla_{\cdot_2}(-\Delta_{\cdot_1,\cdot_2})^{-\frac12}(-\Delta_{\cdot_2})^{-\frac12}(-\Delta_{\cdot_1,\cdot_2})^{\frac12}(-\Delta_{\cdot_2})^{\frac12}f)(\cdot_1,\cdot))(\cdot_2)\big)(x_1,x_2)\Big)\,dx_1\,dx_2$$
$$\le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|(-\Delta_{x_1,x_2})^{\frac12}(-\Delta_{x_2})^{\frac12}f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}, \tag{3-24}$$

where in the second inequality the area function $S_{\mathcal{F}}$ is defined as in Definition 2.2, and the last inequality follows from Hölder's inequality and boundedness of the maximal functions as well as the boundedness of the flag Riesz transforms. Hence we see that (3-8) holds.

By using an estimate similar to that above, we can obtain the estimates in (3-9)–(3-23). We omit the details here since they are straightforward. □

Upper bound for iterated commutators.

Theorem 3.6. *For every $b \in \mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$, $g \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^m)$ and for any $i = 1, 2, \dots, m+n$, $j = 1, \dots, n$, there exits a positive constant C depending only on p, n and m such that*

$$\|[[b, R_i^{(1)}], R_j^{(2)}]_2(g)\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}. \tag{3-25}$$

Proof. Recall that

$$\begin{aligned}[][b, R_i^{(1)}], R_j^{(2)}]_2(g)(x_1, x_2) = b(x_1, x_2)R_i^{(1)} * R_j^{(2)} *_2 g(x_1, x_2) - R_i^{(1)} * (b \cdot R_j^{(2)} *_2 g)(x_1, x_2)\\ - R_j^{(2)} *_2 (b \cdot R_i^{(1)} * g)(x_1, x_2) + R_j^{(2)} *_2 R_i^{(1)} * (b \cdot g)(x_1, x_2).\end{aligned}$$

Hence, for every $f \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^m)$, we have

$$\begin{aligned}\langle f, [[b, R_i^{(1)}], R_j^{(2)}]_2(g)\rangle = \langle f \cdot b, R_i^{(1)} * R_j^{(2)} *_2 g\rangle + \langle R_i^{(1)} * f, b \cdot R_j^{(2)} *_2 g\rangle\\ + \langle R_j^{(2)} *_2 f, b \cdot R_i^{(1)} * g\rangle + \langle R_j^{(2)} *_2 R_i^{(1)} * f, b \cdot g\rangle.\end{aligned}$$

Denote by B, F, G the flag harmonic extensions of the functions b, f, g, respectively, as defined in (3-1). And for each fixed i, j, denote by $(R_i^{(1)} * f)^\sim$, $(R_j^{(2)} *_2 f)^\sim$ and $(R_i^{(1)} * R_j^{(2)} *_2 f)^\sim$ the flag harmonic extensions of $R_i^{(1)} * f$, $R_j^{(2)} *_2 f$ and $R_i^{(1)} * R_j^{(2)} *_2 f$.

Then we write

$$\begin{aligned}&\langle f, [[b, R_i^{(1)}], R_j^{(2)}]_2(g)\rangle\\ &= \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 \big(F\cdot B\cdot(R_i^{(1)}*R_j^{(2)}*_2 g)^\sim + (R_i^{(1)}*f)^\sim\cdot B\cdot(R_j^{(2)}*_2 g)^\sim\\ &\qquad + (R_j^{(2)}*_2 f)^\sim\cdot B\cdot(R_i^{(1)}*g)^\sim + (R_j^{(2)}*_2 R_i^{(1)}*f)^\sim\cdot B\cdot G\big)\, dx_1\, dx_2\, dt_1\, dt_2.\end{aligned} \tag{3-26}$$

We now claim that the right-hand side of (3-26) is bounded by

$$C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}. \tag{3-27}$$

To see this, we compute the derivatives $t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2$ for the integrand in the right-hand side of (3-26). Then we have the following terms:

$$\begin{aligned}\mathcal{C}_1 = \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} \Big(& t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 B\cdot F\cdot(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim + t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 B\cdot(R_i^{(1)} * f)^\sim\cdot(R_j^{(2)} *_2 g)^\sim\\ &+ t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 B\cdot(R_j^{(2)} *_2 f)^\sim\cdot(R_i^{(1)} * g)^\sim\\ &+ t_1\partial_{t_1}^2\, t_2\partial_{t_2}^2 B\cdot(R_j^{(2)} *_2 R_i^{(1)} * f)^\sim\cdot G\Big)\, dx_1\, dx_2\, dt_1\, dt_2;\end{aligned} \tag{3-28}$$

$$\begin{aligned}\mathcal{C}_2=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1\partial^2_{t_1}\,t_2\partial_{t_2}B\cdot\partial_{t_2}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1\partial^2_{t_1}\,t_2\partial_{t_2}B\cdot\partial_{t_2}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1\partial^2_{t_1}\,t_2\partial_{t_2}B\cdot\partial_{t_2}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1\partial^2_{t_1}\,t_2\partial_{t_2}B\cdot\partial_{t_2}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-29}$$

$$\begin{aligned}\mathcal{C}_3=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1\,\partial_{t_1}t_2\,\partial^2_{t_2}B\cdot\partial_{t_1}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1\,\partial_{t_1}t_2\,\partial^2_{t_2}B\cdot\partial_{t_1}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1\,\partial_{t_1}t_2\,\partial^2_{t_2}B\cdot\partial_{t_1}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1\,\partial_{t_1}t_2\,\partial^2_{t_2}B\cdot\partial_{t_1}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-30}$$

$$\begin{aligned}\mathcal{C}_4=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1\,\partial_{t_1}t_2\,\partial_{t_2}B\cdot\partial_{t_1}\partial_{t_2}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1\,\partial_{t_1}t_2\,\partial_{t_2}B\cdot\partial_{t_1}\partial_{t_2}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1\,\partial_{t_1}t_2\,\partial_{t_2}B\cdot\partial_{t_1}\partial_{t_2}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1\,\partial_{t_1}t_2\,\partial_{t_2}B\cdot\partial_{t_1}\partial_{t_2}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-31}$$

$$\begin{aligned}\mathcal{C}_5=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1\,\partial_{t_1}t_2B\cdot\partial_{t_1}\partial^2_{t_2}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1\,\partial_{t_1}t_2B\cdot\partial_{t_1}\partial^2_{t_2}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1\,\partial_{t_1}t_2B\cdot\partial_{t_1}\partial^2_{t_2}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1\,\partial_{t_1}t_2B\cdot\partial_{t_1}\partial^2_{t_2}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-32}$$

$$\begin{aligned}\mathcal{C}_6=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1t_2\,\partial_{t_2}B\cdot\partial^2_{t_1}\partial_{t_2}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1t_2\partial_{t_2}B\cdot\partial^2_{t_1}\partial_{t_2}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1t_2\,\partial_{t_2}B\cdot\partial^2_{t_1}\partial_{t_2}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1t_2\,\partial_{t_2}B\cdot\partial^2_{t_1}\partial_{t_2}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-33}$$

$$\begin{aligned}\mathcal{C}_7=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1t_2\,\partial^2_{t_2}B\cdot\partial^2_{t_1}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1t_2\,\partial^2_{t_2}B\cdot\partial^2_{t_1}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1t_2\,\partial^2_{t_2}B\cdot\partial^2_{t_1}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1t_2\,\partial^2_{t_2}B\cdot\partial^2_{t_1}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-34}$$

$$\begin{aligned}\mathcal{C}_8=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1t_2\,\partial^2_{t_1}B\cdot\partial^2_{t_2}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1t_2\,\partial^2_{t_1}B\cdot\partial^2_{t_2}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1t_2\,\partial^2_{t_1}B\cdot\partial^2_{t_2}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1t_2\,\partial^2_{t_1}B\cdot\partial^2_{t_2}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2;\end{aligned}\tag{3-35}$$

$$\begin{aligned}\mathcal{C}_9=\int_{\mathbb{R}^{n+1}_+\times\mathbb{R}^{m+1}_+} & t_1t_2B\cdot\partial^2_{t_1}\partial^2_{t_2}(F\cdot(R^{(1)}_i*R^{(2)}_j*_2 g)^\sim)+t_1t_2B\cdot\partial^2_{t_1}\partial^2_{t_2}((R^{(1)}_i*f)^\sim\cdot(R^{(2)}_j*_2 g)^\sim)\\ &+t_1t_2B\cdot\partial^2_{t_1}\partial^2_{t_2}((R^{(2)}_j*_2 f)^\sim\cdot(R^{(1)}_i*g)^\sim)\\ &+t_1t_2B\cdot\partial^2_{t_1}\partial^2_{t_2}((R^{(2)}_j*_2 R^{(1)}_i*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2.\end{aligned}\tag{3-36}$$

We first consider $\mathcal{C}_1$. Note that $\partial_{t_2}^2 B = -\Delta_{x_2} B = -\nabla_{x_2}\cdot\nabla_{x_2} B$ and that $\partial_{t_1}^2 B = -\Delta_{x_1,x_2} B = -\nabla_{x_1,x_2}\cdot\nabla_{x_1,x_2} B$. So, integration by parts gives

$$\begin{aligned}\mathcal{C}_1 = \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} & t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_1,x_2}\nabla_{x_2}(F\cdot(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim) \\ & + t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_1,x_2}\nabla_{x_2}((R_i^{(1)} * f)^\sim\cdot(R_j^{(2)} *_2 g)^\sim) \\ & \quad + t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_1,x_2}\nabla_{x_2}((R_j^{(2)} *_2 f)^\sim\cdot(R_i^{(1)} * g)^\sim) \\ & \qquad + t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_1,x_2}\nabla_{x_2}((R_j^{(2)} *_2 R_i^{(1)} * f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2 \\ =:{}& \mathcal{C}_{1,1}+\mathcal{C}_{1,2}+\mathcal{C}_{1,3}+\mathcal{C}_{1,4}.\end{aligned}$$

For the first term, it is clear that

$$\begin{aligned}\mathcal{C}_{1,1} &= \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_1,x_2}\nabla_{x_2}F\cdot(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim\,dx_1\,dx_2\,dt_1\,dt_2 \\ &= \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_1,x_2}F\cdot\nabla_{x_2}(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim\,dx_1\,dx_2\,dt_1\,dt_2 \\ &= \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot\nabla_{x_2}F\cdot\nabla_{x_1,x_2}(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim\,dx_1\,dx_2\,dt_1\,dt_2 \\ &= \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot F\cdot\nabla_{x_1,x_2}\nabla_{x_2}(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim\,dx_1\,dx_2\,dt_1\,dt_2 \\ &=: \mathcal{C}_{1,1,1}+\mathcal{C}_{1,1,2}+\mathcal{C}_{1,1,3}+\mathcal{C}_{1,1,4}.\end{aligned}$$

It is direct that $\mathcal{C}_{1,1,1}$ and $\mathcal{C}_{1,1,4}$ can be handled by using (3-9), and $\mathcal{C}_{1,1,2}$ and $\mathcal{C}_{1,1,3}$ can be handled by using (3-10), which gives

$$\mathcal{C}_{1,1} \le C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}.$$

Symmetrically we obtain the estimate for $\mathcal{C}_{1,4}$, and using similar estimates we can handle $\mathcal{C}_{1,2}$ and $\mathcal{C}_{1,3}$. All these three terms are have the same upper as $\mathcal{C}_{1,1}$ above.

Next, for $\mathcal{C}_2$, note that $\partial_{t_1}^2 B = -\Delta_{x_1,x_2} B = -\nabla_{x_1,x_2}\cdot\nabla_{x_1,x_2} B$. Thus, similar to the term $\mathcal{C}_1$, by integration by parts, we have

$$\begin{aligned}\mathcal{C}_2 = -\int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} & t_1t_2\nabla_{x_1,x_2}\,\partial_{t_2}B\cdot\nabla_{x_1,x_2}\,\partial_{t_2}(F\cdot(R_i^{(1)} * R_j^{(2)} *_2 g)^\sim) \\ & + t_1t_2\nabla_{x_1,x_2}\,\partial_{t_2}B\cdot\nabla_{x_1,x_2}\,\partial_{t_2}((R_i^{(1)} * f)^\sim\cdot(R_j^{(2)} *_2 g)^\sim) \\ & \quad + t_1t_2\nabla_{x_1,x_2}\,\partial_{t_2}B\cdot\nabla_{x_1,x_2}\,\partial_{t_2}((R_j^{(2)} *_2 f)^\sim\cdot(R_i^{(1)} * g)^\sim) \\ & \qquad + t_1t_2\nabla_{x_1,x_2}\,\partial_{t_2}B\cdot\nabla_{x_1,x_2}\,\partial_{t_2}((R_j^{(2)} *_2 R_i^{(1)} * f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2 \\ =:{}& \mathcal{C}_{2,1}+\mathcal{C}_{2,2}+\mathcal{C}_{2,3}+\mathcal{C}_{2,4}.\end{aligned}$$

Again, the upper bounds from the four terms above can be obtained by applying Proposition 3.5, and they are all controlled by

$$C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}.$$

The term $\mathcal{C}_3$ can be handled symmetrically to $\mathcal{C}_2$ and we obtain the same upper bounds.

For the term $\mathcal{C}_4$, by noting that $|\partial_{t_1}\partial_{t_2}B(x_1,x_2,t_1,t_2)|$ is bounded by $|\nabla^{(1)}\nabla^{(2)}B(x_1,x_2,t_1,t_2)|$, we obtain that $\mathcal{C}_4$ is bounded by

$$C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)},$$

where we apply again the upper bounds in Proposition 3.5.

We now turn to the term $\mathcal{C}_9$. We first point out the following equalities:

$$\begin{aligned}
\partial_{t_1}(R_i^{(1)}*R_j^{(2)}*_2 g)^{\sim}(x_1,x_2) &= -c\,\partial_{(x_1,x_2),i}\,(R_j^{(2)}*_2 g)^{\sim}(x_1,x_2),\\
\partial_{t_1}^2(R_i^{(1)}*R_j^{(2)}*_2 g)^{\sim}(x_1,x_2) &= -c\,\partial_{t_1}\partial_{(x_1,x_2),i}\,(R_j^{(2)}*_2 g)^{\sim}(x_1,x_2),\\
\partial_{t_2}(R_i^{(1)}*R_j^{(2)}*_2 g)^{\sim}(x_1,x_2) &= -c\,\partial_{x_{2,j}}\,(R_i^{(1)}*g)^{\sim}(x_1,x_2),\\
\partial_{t_2}^2(R_i^{(1)}*R_j^{(2)}*_2 g)^{\sim}(x_1,x_2) &= -c\,\partial_{t_2}\partial_{x_{2,j}}\,(R_i^{(1)}*g)^{\sim}(x_1,x_2),\\
\partial_{t_1}(R_i^{(1)}*f)^{\sim} &= -c\,\partial_{(x_1,x_2),i}(f)^{\sim},\\
\partial_{t_1}^2(R_i^{(1)}*f)^{\sim} &= -c\,\partial_{t_1}\partial_{(x_1,x_2),i}(f)^{\sim},\\
\partial_{t_2}(R_j^{(2)}*_2 g)^{\sim} &= -c\,\partial_{x_{2,j}}(g)^{\sim},\\
\partial_{t_2}^2(R_j^{(2)}*_2 g)^{\sim} &= -c\,\partial_{t_2}\partial_{x_{2,j}}(g)^{\sim}.
\end{aligned}$$

Then for the term $\mathcal{C}_9$, we get

$$\begin{aligned}
&\partial_{t_1}^2\partial_{t_2}^2(F\cdot(R_i^{(1)}*R_j^{(2)}*_2g)^{\sim}+(R_i^{(1)}*f)^{\sim}\cdot(R_j^{(2)}*_2g)^{\sim}+(R_j^{(2)}*_2f)^{\sim}\cdot(R_i^{(1)}*g)^{\sim}+(R_j^{(2)}*_2R_i^{(1)}*f)^{\sim}\cdot G)\\
&=4\,\partial_{(x_1,x_2),i}\partial_{t_1}\partial_{x_{2,j}}\partial_{t_2}(FG)\\
&\qquad-2\nabla_{x_1,x_2}\,\partial_{x_{2,j}}\partial_{t_2}(\nabla_{x_1,x_2}(R_i^{(1)}*f)^{\sim}\cdot G)-2\nabla_{x_1,x_2}\,\partial_{x_{2,j}}\partial_{t_2}(F\cdot\nabla_{x_1,x_2}(R_i^{(1)}*g)^{\sim})\\
&\qquad+2\nabla_{x_1,x_2}\,\partial_{x_{2,j}}\partial_{t_2}(\nabla_{x_1,x_2}F\cdot(R_i^{(1)}*g)^{\sim})+2\nabla_{x_1,x_2}\,\partial_{x_{2,j}}\partial_{t_2}((R_i^{(1)}*f)^{\sim}\cdot\nabla_{x_1,x_2}G)\\
&\quad-2\,\partial_{(x_1,x_2),i}\partial_{t_1}\nabla_{x_2}(\nabla_{x_2}(R_j^{(2)}*f)^{\sim}\cdot G)\\
&\qquad+\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}\nabla_{x_2}(R_j^{(2)}*_2R_i^{(1)}*f)^{\sim}\cdot G)+\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_2}(R_j^{(2)}*_2f)^{\sim}\cdot\nabla_{x_1,x_2}(R_i^{(1)}*g)^{\sim})\\
&\qquad-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}\nabla_{x_2}(R_j^{(2)}*_2f)^{\sim}\cdot(R_i^{(1)}*g)^{\sim})-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_2}(R_j^{(2)}*_2R_i^{(1)}*f)^{\sim}\cdot\nabla_{x_1,x_2}G)\\
&\quad-2\,\partial_{(x_1,x_2),i}\partial_{t_1}\nabla_{x_2}(F\cdot\nabla_{x_2}(R_j^{(2)}*g)^{\sim})\\
&\qquad+\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}(R_i^{(1)}*f)^{\sim}\cdot\nabla_{x_2}(R_j^{(2)}*_2g)^{\sim})+\nabla_{x_1,x_2}\nabla_{x_2}(F\cdot\nabla_{x_1,x_2}\nabla_{x_2}(R_i^{(1)}*R_j^{(2)}*_2g)^{\sim})\\
&\qquad-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}\nabla_{x_2}(R_j^{(2)}*_2f)^{\sim}\cdot(R_i^{(1)}*g)^{\sim})-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_2}(R_j^{(2)}*_2R_i^{(1)}*f)^{\sim}\cdot\nabla_{x_1,x_2}G)\\
&\quad+2\,\partial_{(x_1,x_2),i}\partial_{t_1}\nabla_{x_2}(\nabla_{x_2}F\cdot(R_j^{(2)}*g)^{\sim})\\
&\qquad-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}\nabla_{x_2}(R_i^{(1)}*f)^{\sim}\cdot(R_j^{(2)}*_2g)^{\sim})-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_2}F\cdot\nabla_{x_1,x_2}(R_i^{(1)}*R_j^{(2)}*_2g)^{\sim})\\
&\qquad+\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}\nabla_{x_2}F\cdot(R_i^{(1)}*R_j^{(2)}*_2g)^{\sim})+\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_2}(R_i^{(1)}*f)^{\sim}\cdot\nabla_{x_1,x_2}(R_j^{(2)}*_2g)^{\sim})\\
&\quad+2\,\partial_{(x_1,x_2),i}\partial_{t_1}\nabla_{x_2}((R_j^{(2)}*f)^{\sim}\cdot\nabla_{x_2}G)\\
&\qquad-\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}(R_i^{(1)}*R_j^{(2)}*_2f)^{\sim}\cdot\nabla_{x_2}G)-\nabla_{x_1,x_2}\nabla_{x_2}((R_j^{(2)}*_2f)^{\sim}\cdot\nabla_{x_1,x_2}\nabla_{x_2}(R_i^{(1)}*g)^{\sim})\\
&\qquad+\nabla_{x_1,x_2}\nabla_{x_2}(\nabla_{x_1,x_2}(R_j^{(2)}*_2f)^{\sim}\cdot\nabla_{x_2}(R_i^{(1)}*g)^{\sim})+\nabla_{x_1,x_2}\nabla_{x_2}((R_i^{(1)}*R_j^{(2)}*_2f)^{\sim}\cdot\nabla_{x_1,x_2}\nabla_{x_2}G).
\end{aligned}$$

Thus, we input the above 25 terms back into the right-hand side of $\mathcal{C}_9$ and obtain the terms as follows:

$$\begin{aligned}\mathcal{C}_9 &= \int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2 B\cdot\partial_{t_1}^2\partial_{t_2}^2\big(F\cdot(R_i^{(1)}*R_j^{(2)}*_2 g)^\sim + (R_i^{(1)}*f)^\sim\cdot(R_j^{(2)}*_2 g)^\sim \\ &\qquad + (R_j^{(2)}*_2 f)^\sim\cdot(R_i^{(1)}*g)^\sim + (R_j^{(2)}*_2 R_i^{(1)}*f)^\sim\cdot G\big)\,dx_1\,dx_2\,dt_1\,dt_2 \\ &= 4\int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\,\partial_{(x_1,x_2),i}\partial_{x_{2,j}}B\cdot\partial_{t_1}\partial_{t_2}(FG)\,dx_1\,dx_2\,dt_1\,dt_2 \\ &\quad -2\int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\nabla_{x_1,x_2}\,\partial_{x_{2,j}}B\cdot\partial_{t_2}(\nabla_{x_1,x_2}(R_i^{(1)}*f)^\sim\cdot G)\,dx_1\,dx_2\,dt_1\,dt_2 \\ &\quad \cdots+\int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1t_2\nabla_{x_1,x_2}\nabla_{x_2}B\cdot((R_i^{(1)}*R_j^{(2)}*_2 f)^\sim\cdot\nabla_{x_1,x_2}\nabla_{x_2}G)\,dx_1\,dx_2\,dt_1\,dt_2 \\ &= \mathcal{C}_{9,1}+\mathcal{C}_{9,2}+\cdots+\mathcal{C}_{9,25},\end{aligned}$$

where we get all these terms from the equality $\partial_{t_1}^2\partial_{t_2}^2(\cdots)$ by integration by parts and taking all the gradients or partial derivatives with respect to x_1, x_2 to the function B. By applying Proposition 3.5 to all these terms, we obtain that they are all controlled by

$$C\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^{p'}(\mathbb{R}^n\times\mathbb{R}^m)}.$$

Next we consider the term $\mathcal{C}_5$, which can be considered as a cross term in between $\mathcal{C}_1$ and $\mathcal{C}_9$. To continue, we write

$$\begin{aligned}&\partial_{t_2}^2(F\cdot(R_i^{(1)}*R_j^{(2)}*_2 g)^\sim + (R_i^{(1)}*f)^\sim\cdot(R_j^{(2)}*_2 g)^\sim + (R_j^{(2)}*_2 f)^\sim\cdot(R_i^{(1)}*g)^\sim + (R_j^{(2)}*_2 R_i^{(1)}*f)^\sim\cdot G) \\ &\quad = \partial_{t_2}^2(F\cdot(R_j^{(2)}*_2(R_i^{(1)}*g))^\sim + (R_j^{(2)}*_2 f)^\sim\cdot(R_i^{(1)}*g)^\sim) \\ &\qquad + \partial_{t_2}^2((R_i^{(1)}*f)^\sim\cdot(R_j^{(2)}*_2 g)^\sim + (R_j^{(2)}*_2(R_i^{(1)}*f))^\sim\cdot G) \\ &\quad = E_1+E_2.\end{aligned}$$

For the term E_1, we write

$$\begin{aligned}E_1 &= -2\,\partial_{x_{2,j}}\partial_{t_2}(F\cdot(R_i^{(1)}*g)^\sim) + \nabla_{x_2}\big(\nabla_{x_2}(R_j^{(2)}*_2 f)^\sim\cdot(R_i^{(1)}*g)^\sim + F\cdot\nabla_{x_2}(R_j^{(2)}*_2 R_i^{(1)}*g)^\sim \\ &\qquad - \nabla_{x_2}F\cdot(R_j^{(2)}*_2 R_i^{(1)}*g)^\sim - (R_j^{(2)}*_2 f)^\sim\cdot\nabla_{x_2}(R_i^{(1)}*g)^\sim\big).\end{aligned}$$

For the term E_2, we write

$$\begin{aligned}E_2 &= -2\,\partial_{x_{2,j}}\partial_{t_2}((R_i^{(1)}*f)^\sim\cdot G) + \nabla_{x_2}\big(\nabla_{x_2}(R_j^{(2)}*_2 R_i^{(1)}*f)^\sim\cdot G + (R_i^{(1)}*f)^\sim\cdot\nabla_{x_2}(R_j^{(2)}*_2 g)^\sim \\ &\qquad - \nabla_{x_2}(R_i^{(1)}*f)^\sim\cdot(R_j^{(2)}*_2 g)^\sim - (R_j^{(2)}*_2 f)^\sim\cdot\nabla_{x_2}G\big).\end{aligned}$$

As a consequence, by substituting the above 10 terms in the right-hand side of the equalities E_1 and E_2 back into the term $\mathcal{C}_5$, we have

$$\begin{aligned}\mathcal{C}_5 &= 2\int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1\,\partial_{t_1}t_2\,\partial_{x_{2,j}}B\cdot\partial_{t_1}\partial_{t_2}(F\cdot(R_i^{(1)}*g)^\sim)\,dx_1\,dx_2\,dt_1\,dt_2 \\ &\quad -\int_{\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{m+1}} t_1\,\partial_{t_1}t_2\nabla_{x_2}B\cdot\partial_{t_1}(\nabla_{x_2}(R_j^{(2)}*_2 f)^\sim\cdot(R_i^{(1)}*g)^\sim)\,dx_1\,dx_2\,dt_1\,dt_2\end{aligned}$$

$$
\begin{aligned}
&- \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} (F \cdot \nabla_{x_2} (R_j^{(2)} *_2 R_i^{(1)} * g)^{\sim})\, dx_1\, dx_2\, dt_1\, dt_2 \\
&+ \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} (\nabla_{x_2} F \cdot (R_j^{(2)} *_2 R_i^{(1)} * g)^{\sim})\, dx_1\, dx_2\, dt_1\, dt_2 \\
&+ \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} ((R_j^{(2)} *_2 f)^{\sim} \cdot \nabla_{x_2} (R_i^{(1)} * g)^{\sim})\, dx_1\, dx_2\, dt_1\, dt_2 \\
&+ 2 \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2\, \partial_{x_{2,j}} B \cdot \partial_{t_1} \partial_{t_2} ((R_i^{(1)} * f)^{\sim} \cdot G)\, dx_1\, dx_2\, dt_1\, dt_2 \\
&- \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} (\nabla_{x_2} (R_j^{(2)} *_2 R_i^{(1)} * f)^{\sim} \cdot G)\, dx_1\, dx_2\, dt_1\, dt_2 \\
&- \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} ((R_i^{(1)} * f)^{\sim} \cdot \nabla_{x_2} (R_j^{(2)} *_2 g)^{\sim})\, dx_1\, dx_2\, dt_1\, dt_2 \\
&+ \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} (\nabla_{x_2} (R_i^{(1)} * f)^{\sim} \cdot (R_j^{(2)} *_2 g)^{\sim})\, dx_1\, dx_2\, dt_1\, dt_2 \\
&+ \int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1\, \partial_{t_1} t_2 \nabla_{x_2} B \cdot \partial_{t_1} ((R_j^{(2)} *_2 f)^{\sim} \cdot \nabla_{x_2} G)\, dx_1\, dx_2\, dt_1\, dt_2 \\
&=: \mathcal{C}_{5,1} + \cdots + \mathcal{C}_{5,10}.
\end{aligned}
$$

By applying Proposition 3.5 to these terms, we obtain that they are all controlled by

$$C \|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)} \|g\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m)} \|f\|_{L^{p'}(\mathbb{R}^n \times \mathbb{R}^m)}.$$

The estimates for the term $\mathcal{C}_6$ can be handled symmetrically, and we get the same upper bound for $\mathcal{C}_6$ as that for $\mathcal{C}_5$ above.

For the term $\mathcal{C}_7$, first note that $\partial_{t_2}^2 B = -\Delta_{x_2} B = -\nabla_{x_2} \cdot \nabla_{x_2} B$. Hence we can write

$$
\begin{aligned}
\mathcal{C}_7 = -\int_{\mathbb{R}^{n+1}_+ \times \mathbb{R}^{m+1}_+} t_1 t_2 \nabla_{x_2} B \cdot \nabla_{x_2}\, \partial_{t_1}^2 (F \cdot (R_i^{(1)} * R_j^{(2)} *_2 g)^{\sim} + (R_i^{(1)} * f)^{\sim} \cdot (R_j^{(2)} *_2 g)^{\sim} \\
+ (R_j^{(2)} *_2 f)^{\sim} \cdot (R_i^{(1)} * g)^{\sim} + (R_j^{(2)} *_2 R_i^{(1)} * f)^{\sim} \cdot G)\, dx_1\, dx_2\, dt_1\, dt_2.
\end{aligned}
$$

Similar to the calculation in the terms E_1 and E_2 in the estimate of $\mathcal{C}_5$, we can now decompose

$$\partial_{t_1}^2 (F \cdot (R_i^{(1)} * R_j^{(2)} *_2 g)^{\sim} + (R_i^{(1)} * f)^{\sim} \cdot (R_j^{(2)} *_2 g)^{\sim} + (R_j^{(2)} *_2 f)^{\sim} \cdot (R_i^{(1)} * g)^{\sim} + (R_j^{(2)} *_2 R_i^{(1)} * f)^{\sim} \cdot G)$$

into 10 terms, which further gives

$$C_7 = \mathcal{C}_{7,1} + \cdots + \mathcal{C}_{7,10}.$$

Then by applying Proposition 3.5 to these terms, we obtain that they are all controlled by

$$C \|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)} \|g\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m)} \|f\|_{L^{p'}(\mathbb{R}^n \times \mathbb{R}^m)}.$$

The estimates for the term $\mathcal{C}_8$ can be handled symmetrically, and we get the same upper bound for $\mathcal{C}_7$ above. $\square$

4. Upper bound of the big commutator $[b, R_{j,k}]$

We derive a general upper bound result for commutators of any flag singular integral. The proof is based on the $A_{\mathcal{F},p}$ weighted estimate of flag singular integral operators and a Cauchy integral trick that goes back to the work of Coifman, Rochberg, and Weiss [Coifman et al. 1976]. Roughly speaking, this technique allows one to bootstrap the weighted estimate for an arbitrary linear operator to that of its commutators of any order. This is the first time this idea is explored in the multiparameter flag setting. In fact, although not needed for our upper bound proof, we demonstrate the bootstrapping result in the general higher-order, two-weight setting.

A_p weight and little bmo in the flag setting. To begin with, we define the Muckenhoupt A_p weights in the flag setting, which consists of positive, locally integrable functions w satisfying

$$[w]_{A_{\mathcal{F},p}} := \sup_{R\in\mathcal{R}_{\mathcal{F}}} \left(\frac{1}{|R|}\int_R w(x,y)\,dx\,dy\right)\left(\frac{1}{|R|}\int_R w(x,y)^{1-p'}\,dx\,dy\right)^{p-1} < \infty, \quad 1<p<\infty, \tag{4-1}$$

where p' denotes the Hölder conjugate of p. The following result of [Wu 2014] provides a way of approaching the $A_{\mathcal{F},p}$ weights via the classical weights:

$$A_{\mathcal{F},p} = A_p \cap A_p^{(2)} \quad \text{for all } 1<p<\infty, \tag{4-2}$$

where A_p is the classical Muckenhoupt A_p class of weights on $\mathbb{R}^{n+m}$, and $A_p^{(2)}$ consists of weights $w(x,y)$ such that $w(x,\cdot)\in A_p$ with uniformly bounded characteristics for a.e. fixed $x\in\mathbb{R}^n$.

We first show that a similar relation holds true for $\mathrm{bmo}_{\mathcal{F}}$, which will be a useful tool for us in the study of this space.

Lemma 4.1. *Let* $\mathrm{BMO}(\mathbb{R}^{n+m})$ *denote the classical John–Nirenberg BMO space on* $\mathbb{R}^{n+m}$, *and* $\mathrm{BMO}^{(2)}(\mathbb{R}^m)$ *be the space consisting of functions* $f(x,y)$ *such that* $f(x,\cdot)\in\mathrm{BMO}(\mathbb{R}^m)$ *for a.e. fixed* $x\in\mathbb{R}^n$ *with uniformly bounded norm. There holds*

$$\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m}) = \mathrm{BMO}(\mathbb{R}^{n+m})\cap\mathrm{BMO}^{(2)}(\mathbb{R}^m)$$

with comparable norms.

Proof. The inclusion

$$\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m}) \subset \mathrm{BMO}(\mathbb{R}^{n+m})\cap\mathrm{BMO}^{(2)}(\mathbb{R}^m)$$

can be easily verified. Indeed, the inclusion $\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})\subset\mathrm{BMO}(\mathbb{R}^{n+m})$ is obvious from the definition. Now fix $x\in\mathbb{R}^n$. For any cube $J\subset\mathbb{R}^m$, one can find a sequence of cubes $I_k\subset\mathbb{R}^n$ such that $\ell(I_k)\le\ell(J)$ and I_k shrinks to the point $\{x\}$ as $k\to\infty$. The containment thus follows from the Lebesgue differentiation theorem.

The other inclusion ("$\supset$") of the lemma follows from Proposition 4.2 below, which establishes the exp-log connection between $A_{\mathcal{F},p}$ weights and $\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})$, much as in the one-parameter and the product settings. □

Proposition 4.2. *Suppose w is a weight and $1 < p < \infty$. We have*

(i) *if $w \in A_{\mathcal{F},p}$, then $\log w \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})$;*

(ii) *if $\log w \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})$, then $w^{\eta} \in A_{\mathcal{F},p}$ for sufficiently small $\eta > 0$.*

Proof. One observes directly from the definition that

$$A_{\mathcal{F},p} \subset A_{\mathcal{F},q} \quad \text{for all } 1 < p \le q < \infty,$$

and

$$w \in A_{\mathcal{F},p} \Longleftrightarrow w^{1-p'} \in A_{\mathcal{F},p'} \quad \text{for all } 1 < p < \infty.$$

Therefore, it suffices to prove the case $p = 2$.

We first prove (i). Suppose $w \in A_{\mathcal{F},2}$ and let $\gamma = \log w$. Then, for any $R \in \mathcal{R}_{\mathcal{F}}$ the $A_{\mathcal{F},2}$ condition implies

$$\left(\frac{1}{|R|}\int_R e^{\gamma(x,y)-\langle\gamma\rangle_R}\,dx\,dy\right)\left(\frac{1}{|R|}\int_R e^{\langle\gamma\rangle_R-\gamma(x,y)}\,dx\,dy\right) \le [w]_{A_{\mathcal{F},2}} < \infty.$$

By Jensen's inequality we have each of the factors above is at least 1 and at most $[w]_{A_{\mathcal{F},2}}$. Therefore, the inequality

$$\frac{1}{|R|}\int_R e^{|\gamma(x,y)-\langle\gamma\rangle_R|}\,dx\,dy \le 2[w]_{A_{\mathcal{F},2}}$$

holds, which, using the trivial estimate $t \le e^t$, implies

$$\frac{1}{|R|}\int_R |\gamma(x,y)-\langle\gamma\rangle_R|\,dx\,dy \le 2[w]_{A_{\mathcal{F},2}}.$$

Hence, $\gamma \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})$.

We now prove (ii). Let $\gamma = \log w \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})$; it follows from Lemma 4.1 that $\gamma \in \mathrm{BMO}(\mathbb{R}^{n+m})$ and $\gamma \in \mathrm{BMO}^{(2)}(\mathbb{R}^m)$. According to the classical exp-log connection between BMO and A_2, there hold for sufficiently small $\eta > 0$

$$\begin{aligned} &e^{\eta\gamma(\cdot,\cdot)} \in A_2(\mathbb{R}^{n+m}),\\ &e^{\eta\gamma(x,\cdot)} \in A_2(\mathbb{R}^m) \quad \text{uniformly in } x \in \mathbb{R}^n. \end{aligned}$$

Hence, (4-2) implies $e^{\eta\gamma} \in A_{\mathcal{F},2}$ for sufficiently small $\eta > 0$, which completes the proof. □

Upper bound of the commutator. Given an operator T, define its k-th order commutator as

$$C^k_{\vec{b}}(T) := [b_k, [b_{k-1}, \dots, [b_1, T]\cdots]],$$

where each b_j is a function on $\mathbb{R}^n \times \mathbb{R}^m$ for all $1 \le j \le k$.

Theorem 4.3. *Let ν be a fixed weight on $\mathbb{R}^n \times \mathbb{R}^m$, $1 < p < \infty$, and T be a linear operator satisfying*

$$\|T\|_{L^p(\mu)\to L^p(\lambda)} \le C_{n,m,p,T}([\mu]_{A_{\mathcal{F},p}}, [\lambda]_{A_{\mathcal{F},p}}),$$

where $C_{n,m,p,T}(\cdot\,,\cdot)$ *is an increasing function of both components, with* $\mu, \lambda \in A_{\mathcal{F},p}$ *and* $\mu/\lambda = \nu^p$. *For* $k \geq 1$, *let* $b_j \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$, $1 \leq j \leq k$; *then there holds*

$$\|C_{\vec{b}}^k(T)\|_{L^p(\mu)\to L^p(\lambda)} \leq C_{n,m,p,k,T}([\mu]_{A_{\mathcal{F},p}}, [\lambda]_{A_{\mathcal{F},p}}) \prod_{j=1}^{k} \|b_j\|_{\mathrm{bmo}_{\mathcal{F}}}.$$

Assuming Theorem 4.3, in order to derive an (even unweighted) upper estimate for commutator of operator T, it suffices to know the corresponding weighted estimate for T itself. When T is a flag singular integral operator (which includes the flag Riesz transform $R_{j,k}$), such a result was obtained by Han, Lin and Wu [Han et al. 2016b].

Definition 4.4. A flag singular integral $T_{\mathcal{F}} : f \mapsto \mathcal{K} * f$ is defined via a flag kernel $\mathcal{K}$ on $\mathbb{R}^n \times \mathbb{R}^m$, which is a distribution on $\mathbb{R}^{n+m}$ that coincides with a C^∞ function away from the coordinate subspace $\{(0, y)\} \subset \mathbb{R}^{n+m}$ and satisfies:

(i) (differential inequalities) For each $\alpha = (\alpha_1, \dots, \alpha_n)$, $\beta = (\beta_1, \dots, \beta_n)$

$$|\partial_x^\alpha \partial_y^\beta \mathcal{K}(x, y)| \lesssim |x|^{-n-|\alpha|}(|x| + |y|)^{-m-|\beta|}$$

for all $(x, y) \in \mathbb{R}^{n+m}$ with $|x| \neq 0$.

(ii) (cancellation conditions)

$$\left|\int_{\mathbb{R}^m} \partial_x^\alpha \mathcal{K}(x, y)\psi_1(\delta y)\, dy\right| \leq C_\alpha |x|^{-n-|\alpha|}$$

for every multi-index α and for every normalized bump function ψ_1 on $\mathbb{R}^m$ and every $\delta > 0$;

$$\left|\int_{\mathbb{R}^n} \partial_y^\beta \mathcal{K}(x, y)\psi_2(\delta y)\, dy\right| \leq C_\beta |y|^{-m-|\beta|}$$

for every multi-index β and for every normalized bump function ψ_2 on $\mathbb{R}^n$ and every $\delta > 0$;

$$\left|\int_{\mathbb{R}^{n+m}} \mathcal{K}(x, y)\psi_3(\delta_1 x, \delta_2 y)\, dx\, dy\right| \leq C$$

for every normalized bump function ψ_3 on $\mathbb{R}^{n+m}$ and every $\delta_1, \delta_2 > 0$.

Theorem 4.5 [Han et al. 2016b, Remark 1.4]. *Let* $1 < p < \infty$ *and* $w \in A_{\mathcal{F},p}(\mathbb{R}^{n+m})$; *there holds*

$$\|T_{\mathcal{F}}(f)\|_{L_w^p(\mathbb{R}^{n+m})} \leq C_p \|f\|_{L_w^p(\mathbb{R}^{n+m})} \quad \textit{for all } f \in L_w^p(\mathbb{R}^{n+m}).$$

Applying Theorem 4.3 (with the choice $\mu = \lambda = w$) together with Theorem 4.5, one obtains immediately the following.

Corollary 4.6. *Let* $w \in A_{\mathcal{F},p}$, $1 < p < \infty$, *and* T *be a flag singular integral operator as defined above. For any* $k \geq 1$, $\vec{b} = (b_1, \dots, b_k)$ *where* $b_j \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$, $j = 1, \dots, k$, *there holds*

$$\|C_{\vec{b}}^k(T)\|_{L^p(w)\to L^p(w)} \leq C_{n,m,p,k,w,T} \prod_{j=1}^{k} \|b_j\|_{\mathrm{bmo}_{\mathcal{F}}}.$$

Obviously, the result above in the first-order unweighted case is precisely the desired upper bound estimate in Theorem 1.5.

The core of the proof of Theorem 4.3 lies in a complex function representation of the commutators and the Cauchy integral formula. This method has been widely used to obtain upper estimates for linear and multilinear commutators in various settings; see [Chung et al. 2012; Coifman et al. 1976; Hytönen 2016; Bényi et al. 2017; Kunwar and Ou 2017] for examples. The main new challenge in our problem is the unique structure of the little flag BMO space and flag weights, which for instance doesn't seem to fall into the category of spaces recently studied in [Bényi et al. 2017].

Proof of Theorem 4.3. Observe that

$$C^k_{\vec{b}}(T) = \partial_{z_1} \cdots \partial_{z_k} F(\vec{0}), \quad F(\vec{z}) := e^{\sum_{j=1}^k b_1 z_1} T e^{-\sum_{j=1}^k b_j z_j},$$

which generalizes a classical formula representing higher-order commutators. We remark that when all the symbol functions b_j are the same, one can work instead with a simpler formula using single variable complex functions and their k-th order derivatives. According to the Cauchy integral formula on polydiscs,

$$C^k_{\vec{b}}(T) = \frac{1}{(2\pi i)^k} \oint \cdots \oint \frac{F(\vec{z})\, dz_1 \cdots dz_k}{z_1^2 \cdots z_k^2},$$

where each integral is over any closed path around the origin in the corresponding variable. For fixed $(\delta_1, \dots, \delta_k)$ which will be determined later, there holds by the Minkowski inequality

$$\begin{aligned} &\|C^k_{\vec{b}}(T)\|_{L^p(\mu)\to L^p(\lambda)} \\ &\quad\le \frac{1}{(2\pi)^k} \oint_{|z_1|=\delta_1} \cdots \oint_{|z_k|=\delta_k} \|T\|_{L^p(e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\mu)\to L^p(e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\lambda)} \frac{|dz_1|\cdots|dz_k|}{\delta_1^2\cdots\delta_k^2} \\ &\quad\le \frac{1}{(2\pi)^k} \oint_{|z_1|=\delta_1} \cdots \oint_{|z_k|=\delta_k} C_{n,m,p,T}([e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\mu]_{A_{\mathcal{F},p}}, [e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\lambda]_{A_{\mathcal{F},p}}) \frac{|dz_1|\cdots|dz_k|}{\delta_1^2\cdots\delta_k^2}, \end{aligned}$$

where we have used the fact that $(e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\mu, e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\lambda)$ is a pair of weights satisfying

$$\frac{e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\mu}{e^{p\operatorname{Re}(\sum_{j=1}^k b_j z_j)}\lambda} = \frac{\mu}{\lambda} = \nu^p.$$

Now we choose $\{\delta_j\}$ according to Lemma 4.7 below, which is the key ingredient of the proof concerning the relation between $A_{\mathcal{F},p}$ weights and little flag BMO functions. Let

$$\delta_1 := \frac{\epsilon_{n,m,p}}{\max((\mu)_{A_{\mathcal{F},p}}, (\lambda)_{A_{\mathcal{F},p}})\|b_1\|_{\mathrm{bmo}_{\mathcal{F}}}},$$

where for any $w \in A_{\mathcal{F},p}$

$$(w)_{A_{\mathcal{F},p}} := \max([w]_{A_{\mathcal{F},p}}, [\sigma]_{A_{\mathcal{F},p'}}). \tag{4-3}$$

Here we have used the notation $\sigma := w^{1-p'}$ to denote the dual weight of w, and the relevant property of $(w)_{A_{\mathcal{F},p}}$ to us is that

$$(w)_{A_{\mathcal{F},p}} = \max([w]_{A_{\mathcal{F},p}}, [w]^{p'-1}_{A_{\mathcal{F},p}}) = [w]^{\max(1,p'-1)}_{A_{\mathcal{F},p}}.$$

Recursively, for any $j \geq 2$, choose

$$\delta_j := \frac{\epsilon_{n,m,p}}{\sup_{\{z_t\}:\,|z_1|=\delta_1,\dots,|z_{j-1}|=\delta_{j-1}} \max((e^{p\,\mathrm{Re}\,(\sum_{t=1}^{j-1} b_t z_t)}\mu)_{A_{\mathcal{F},p}}, (e^{p\,\mathrm{Re}\,(\sum_{t=1}^{j-1} b_t z_t)}\lambda)_{A_{\mathcal{F},p}})\|b_j\|_{\mathrm{bmo}_{\mathcal{F}}}}.$$

Then applying Lemma 4.7 iteratively shows that

$$[e^{p\,\mathrm{Re}\,(\sum_{j=1}^{k} b_j z_j)}\mu]_{A_{\mathcal{F},p}} \leq C_{n,m,p}[e^{p\,\mathrm{Re}\,(\sum_{j=1}^{k-1} b_j z_j)}\mu]_{A_{\mathcal{F},p}} \leq \cdots \leq C^k_{n,m,p}[\mu]_{A_{\mathcal{F},p}},$$

and similarly

$$[e^{p\,\mathrm{Re}\,(\sum_{j=1}^{k} b_j z_j)}\lambda]_{A_{\mathcal{F},p}} \leq C^k_{n,m,p}[\lambda]_{A_{\mathcal{F},p}},$$

which in turn via the monotonicity of $C_{n,m,p,T}(\,\cdot\,,\,\cdot\,)$ leads to

$$C_{n,m,p,T}([e^{p\,\mathrm{Re}\,(\sum_{j=1}^{k} b_j z_j)}\mu]_{A_{\mathcal{F},p}}, [e^{p\,\mathrm{Re}\,(\sum_{j=1}^{k} b_j z_j)}\lambda]_{A_{\mathcal{F},p}}) \leq C'_{n,m,p,k,T}([\mu]_{A_{\mathcal{F},p}}, [\lambda]_{A_{\mathcal{F},p}}).$$

Therefore,

$$\begin{aligned}\|C^k_{\vec{b}}(T)\|_{L^p(\mu)\to L^p(\lambda)} &\leq \frac{1}{\delta_1\cdots\delta_k} C'_{n,m,p,k,T}([\mu]_{A_{\mathcal{F},p}}, [\lambda]_{A_{\mathcal{F},p}}) \\ &\leq C_{n,m,p,k,T}([\mu]_{A_{\mathcal{F},p}}, [\lambda]_{A_{\mathcal{F},p}}) \prod_{j=1}^{k} \|b_j\|_{\mathrm{bmo}_{\mathcal{F}}}.\end{aligned}$$ □

Lemma 4.7. *Let $w \in A_{\mathcal{F},p}$, $1 < p < \infty$, and $b \in \mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$. There are constants $\epsilon_{n,m,p}, C_{n,m,p} > 0$ such that*

$$[e^{\mathrm{Re}(bz)}w]_{A_{\mathcal{F},p}} \leq C_{n,m,p}[w]_{A_{\mathcal{F},p}}$$

whenever $z \in \mathbb{C}$ satisfies

$$|z| \leq \frac{\epsilon_{n,m,p}}{\|b\|_{\mathrm{bmo}_{\mathcal{F}}}(w)_{A_{\mathcal{F},p}}},$$

where $(w)_{A_{\mathcal{F},p}}$ is defined as in (4-3).

Proof. This estimate is a consequence of (4-2), Lemma 4.1 and a one-parameter version proven by Hytönen [2016], which states that for any $w \in A_p$, the classical Muckenhoupt A_p class on $\mathbb{R}^d$, $1 < p < \infty$, there exist $\epsilon_{d,p}, C_{d,p} > 0$ such that

$$[e^{\mathrm{Re}(bz)}w]_{A_p} \leq C_{d,p}[w]_{A_p}$$

for all $z \in \mathbb{C}$ with

$$|z| \leq \frac{\epsilon_{n,p}}{\|b\|_{\mathrm{BMO}}(w)_{A_p}}.$$

To see this, by (4-2) and Lemma 4.1, given $w \in A_{\mathcal{F},p}$ and $b \in \mathrm{bmo}_{\mathcal{F}}$, there hold $w \in A_p \cap A^{(2)}_p$ and $b \in \mathrm{BMO}(\mathbb{R}^{n+m}) \cap \mathrm{BMO}^{(2)}(\mathbb{R}^m)$. Hence, taking $\epsilon_{n,m,p} > 0$ sufficiently small, for all $z \in \mathbb{C}$ satisfying

$$|z| \leq \frac{\epsilon_{n,m,p}}{\|b\|_{\mathrm{bmo}_{\mathcal{F}}}(w)_{A_{\mathcal{F},p}}},$$

one has

$$[e^{\mathrm{Re}(bz)}w]_{A_p} \leq C_{n+m,p}[w]_{A_p} \leq C_{n,m,p}[w]_{A_{\mathcal{F},p}}$$

and

$$[e^{\mathrm{Re}(b(x,\cdot)z)}w(x,\cdot)]_{A_p} \le C_{m,p}[w(x,\cdot)]_{A_p} \le C_{n,m,p}[w]_{A_{\mathcal{F},p}} \quad \text{a.e. } x \in \mathbb{R}^n,$$

by observing that

$$\|b\|_{\mathrm{bmo}_{\mathcal{F}}} \gtrsim \max\bigl(\|b\|_{\mathrm{BMO}(\mathbb{R}^{n+m})}, \sup_{x\in\mathbb{R}^n} \|b(x,\cdot)\|_{\mathrm{BMO}^{(2)}(\mathbb{R}^m)}\bigr)$$

and that

$$(w)_{A_{\mathcal{F},p}} \gtrsim \max([w]_{A_p}, \sup_{x\in\mathbb{R}^n}[w(x,\cdot)]_{A_p}). \qquad \square$$

5. Applications: div-curl lemmas in the flag setting

Let $E^{(1)}$ be a vector field on $\mathbb{R}^{n+m}$ taking the values in $\mathbb{R}^{n+m}$, and let $E^{(2)}$ be a vector field on $\mathbb{R}^m$ taking the values in $\mathbb{R}^m$. Now let $\mathcal{M}_{n+m,m}$ denote the set of all $(n+m)\times m$ matrices. We now consider the following version of vector fields on $\mathbb{R}^n \times \mathbb{R}^m$ taking the values in $\mathcal{M}_{n+m,m}$, associated with the flag structure:

$$E = E^{(1)} *_2 E^{(2)} := \begin{bmatrix} E_1^{(1)} *_2 E_1^{(2)} & \dots & E_1^{(1)} *_2 E_m^{(2)} \\ \vdots & \cdots & \vdots \\ E_{n+m}^{(1)} *_2 E_1^{(2)} & \dots & E_{n+m}^{(1)} *_2 E_m^{(2)} \end{bmatrix}, \tag{5-1}$$

where

$$E_j^{(1)} *_2 E_k^{(2)}(x,y) = \int_{\mathbb{R}^m} E_j^{(1)}(x, y-z)E_k^{(2)}(z)\,dz.$$

Next we consider the following L^p space via projections. Suppose $1 < p < \infty$. We define $L^p_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m; \mathcal{M}_{n+m,m})$ to be the set of vector fields E in $L^p(\mathbb{R}^n\times\mathbb{R}^m; \mathcal{M}_{n+m,m})$ such that there exist $r_1, r_2 \in (1,\infty)$ with $1/r_1 + 1/r_2 = 1/p + 1$, $E^{(1)} \in L^{r_1}(\mathbb{R}^{n+m}; \mathbb{R}^{n+m})$, $E^{(2)} \in L^{r_2}(\mathbb{R}^m; \mathbb{R}^m)$ and that $E = E^{(1)} *_2 E^{(2)}$; moreover,

$$\|E\|_{L^p_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})} := \inf \|E^{(1)}\|_{L^{r_1}(\mathbb{R}^{n+m};\mathbb{R}^{n+m})}\|E^{(2)}\|_{L^{r_2}(\mathbb{R}^m;\mathbb{R}^m)},$$

where the infimum is taken over all possible $r_1, r_2 \in (1,\infty)$, $E^{(1)} \in L^{r_1}(\mathbb{R}^{n+m}; \mathbb{R}^{n+m})$, $E^{(2)} \in L^{r_2}(\mathbb{R}^m; \mathbb{R}^m)$.

Given two matrices $A, B \in \mathcal{M}_{n+m,m}$, we define the "dot product" between A and B by

$$A \cdot B = \sum_{j=1}^{n+m}\sum_{k=1}^{m} A_{j,k}B_{j,k}.$$

We point out that this is the Hilbert–Schmidt inner product for two matrices and more generally this is referred to as the Schur product of two matrices.

Proof of Theorem 1.6. Note that B is a vector field on $\mathbb{R}^n\times\mathbb{R}^m$ taking the values in $\mathcal{M}_{n+m,m}$, associated with the flag structure (5-1). Then there exist certain vector fields $B^{(1)}$ on $\mathbb{R}^{n+m}$ taking the values in $\mathbb{R}^{n+m}$ and $B^{(2)}$ on $\mathbb{R}^m$ taking the values in $\mathbb{R}^m$ such that $B = B^{(1)} *_2 B^{(2)}$ and that

$$\|B\|_{L^q_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})} \approx \inf \|B^{(1)}\|_{L^{q_1}(\mathbb{R}^{n+m};\mathbb{R}^{n+m})}\|B^{(2)}\|_{L^{q_2}(\mathbb{R}^m;\mathbb{R}^m)}$$

with $1/q_1 + 1/q_2 = 1/q + 1$.

Thus, $\mathrm{curl}_{(x,y)} B^{(1)} = 0$ implies that there exists $\phi^{(1)} \in L^q(\mathbb{R}^{n+m})$ such that

$$B^{(1)} = (R_1^{(1)}\phi^{(1)}, \dots, R_{n+m}^{(1)}\phi^{(1)})$$

with $\|B^{(1)}\|_{L^{q_1}(\mathbb{R}^{n+m};\mathbb{R}^{n+m})} \approx \|\phi^{(1)}\|_{L^{q_1}(\mathbb{R}^{n+m})}$. Again, $\mathrm{curl}_y B^{(2)} = 0$ implies that there exists $\phi^{(2)} \in L^{q_2}(\mathbb{R}^{n+m})$ such that

$$B^{(2)} = (R_1^{(2)}\phi^{(2)}, \dots, R_m^{(2)}\phi^{(2)})$$

with $\|B^{(2)}\|_{L^{q_2}(\mathbb{R}^m;\mathbb{R}^m)} \approx \|\phi^{(2)}\|_{L^{q_2}(\mathbb{R}^m)}$. As a consequence we get that the matrix B has elements

$$B_{j,k} = R_{j,k} * \phi, \quad j = 1, \dots, n+m, \; k = 1, \dots, m,$$

where $\phi = \phi^{(1)} *_2 \phi^{(2)}$ and $\|B\|_{L^q_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})} \approx \|\phi\|_{L^q(\mathbb{R}^{n+m})}$.

Similarly, note that E is a vector field on $\mathbb{R}^n \times \mathbb{R}^m$ taking the values in $\mathcal{M}_{n+m,m}$, associated with the flag structure (5-1). Then there exist certain vector fields $E^{(1)}$ on $\mathbb{R}^{n+m}$ taking the values in $\mathbb{R}^{n+m}$ and $E^{(2)}$ on $\mathbb{R}^m$ taking the values in $\mathbb{R}^m$ such that $E = E^{(1)} *_2 E^{(2)}$ and that

$$\|E\|_{L^p_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})} \approx \inf \|E^{(1)}\|_{L^{p_1}(\mathbb{R}^{n+m};\mathbb{R}^{n+m})}\|E^{(2)}\|_{L^{p_2}(\mathbb{R}^m;\mathbb{R}^m)}$$

with $1/p_1 + 1/p_2 = 1/p + 1$.

Thus, the conditions $\mathrm{div}_{(x,y)} E^{(1)} = 0$ and $\mathrm{div}_y E^{(2)} = 0$ imply

$$\sum_{j=1}^{n+m} R_j^{(1)} * E_j^{(1)}(x, y) = 0 \quad \text{and} \quad \sum_{k=1}^{m} R_k^{(2)} *_2 E_k^{(2)}(y) = 0.$$

Hence, we get

$$\sum_{j=1}^{n+m} R_j^{(1)} * E_{j,k}(x, y) = 0 \quad \text{and} \quad \sum_{k=1}^{m} R_k^{(2)} *_2 E_{j,k}(x, y) = 0.$$

With these facts, we have

$$\begin{aligned} E(x,y)\cdot B(x,y) &= \sum_{j=1}^{n+m}\sum_{k=1}^{m} E_{j,k}(x,y)B_{j,k}(x,y) = \sum_{j=1}^{n+m}\sum_{k=1}^{m} E_{j,k}(x,y)R_{j,k}*\phi(x,y) \\ &= \sum_{j=1}^{n+m}\sum_{k=1}^{m} \big\{E_{j,k}(x,y)R_{j,k}*\phi(x,y) + R_j^{(1)} * E_{j,k}(x,y)R_k^{(2)} *_2 \phi(x,y) \\ &\qquad\qquad + R_k^{(2)} *_2 E_{j,k}(x,y)R_j^{(1)} * \phi(x,y) + R_{j,k} * E_{j,k}(x,y)\phi(x,y)\big\}. \end{aligned}$$

Now testing this equality over all functions in the flag BMO space, i.e., for every $b \in \mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$, and then unravelling the expression with the Riesz transforms we see that

$$\int_{\mathbb{R}^n\times\mathbb{R}^m} E(x,y)\cdot B(x,y)\, b(x,y)\,dx\,dy = \sum_{j=1}^{n+m}\sum_{k=1}^{m}\int_{\mathbb{R}^n\times\mathbb{R}^m} [[b, R_j^{(1)}], R_k^{(2)}]_2(E_{j,k})(x,y)\phi(x,y)\,dx\,dy.$$

Then based on Theorem 1.3, since $b \in \mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n \times \mathbb{R}^m)$ we have that each of the above commutators is a bounded operator on $L^p(\mathbb{R}^n \times \mathbb{R}^m)$ with norm controlled by the norm of b, i.e., $\|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}$.

As a consequence, we get

$$\left|\int_{\mathbb{R}^n\times\mathbb{R}^m} E(x,y)\cdot B(x,y)\,b(x,y)\,dx\,dy\right| \lesssim \|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|E\|_{L^p_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})}\|\phi\|_{L^q(\mathbb{R}^{n+m})}$$
$$\lesssim \|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|E\|_{L^p_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})}\|B\|_{L^q_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})}.$$

Then from the duality of $H^1_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$ with $\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$, we obtain

$$\|E\cdot B\|_{H^1_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)} \lesssim \|b\|_{\mathrm{BMO}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|E\|_{L^p_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})}\|B\|_{L^q_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m;\mathcal{M}_{n+m,m})}. \qquad \square$$

Proof of Theorem 1.7. Suppose that E, B are vector fields on $\mathbb{R}^n\times\mathbb{R}^m$ taking values in $\mathbb{R}^{n+m}$. Moreover, suppose $E\in L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})$ and $B\in L^q(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})$ satisfy

$$\mathrm{div}_{(x,y)}\,E(x,y)=0 \quad\text{and}\quad \mathrm{curl}_{(x,y)}\,B(x,y)=0$$

and

$$\mathrm{div}_y\,E(x,y)=0 \quad\text{and}\quad \mathrm{curl}_y\,B(x,y)=0 \quad\text{for all } x\in\mathbb{R}^n.$$

We now define the projection operator $\mathcal{P}$ as

$$\mathcal{P}E=\left(E_1+R_1^{(1)}\left(\sum_{k=1}^{n+m}R_k^{(1)}E_k\right),\ldots,E_{n+m}+R_{n+m}^{(1)}\left(\sum_{k=1}^{n+m}R_k^{(1)}E_k\right)\right).$$

Then by definition, it is direct that

$$\mathrm{div}_{(x,y)}\,\mathcal{P}E=0$$

since

$$\sum_{j=1}^{n+m}R_j^{(1)}\left(E_j+R_j^{(1)}\left(\sum_{k=1}^{n+m}R_k^{(1)}E_k\right)\right)=0. \tag{5-2}$$

Moreover, we also have $\mathcal{P}\circ\mathcal{P}E=\mathcal{P}E$. Next, we point out that applying $[b,\mathcal{P}]$ to the vector field E, we can get that the j-th component is given by

$$\sum_{k=1}^{n+m}[b,R_j^{(1)}R_k^{(1)}](E_k).$$

Suppose now $b\in\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$. Then from Lemma 4.1 we know

$$\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^{n+m})=\mathrm{BMO}(\mathbb{R}^{n+m})\cap\mathrm{BMO}^{(2)}(\mathbb{R}^m)$$

with comparable norms. Hence, we have $b\in\mathrm{BMO}(\mathbb{R}^{n+m})$ with

$$\|b\|_{\mathrm{BMO}(\mathbb{R}^{n+m})}\lesssim\|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}.$$

With all these observations, an application of the Coifman, Rochberg and Weiss theorem demonstrates that $[b,\mathcal{P}](E)$ is bounded on $L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})$ with

$$\|[b,\mathcal{P}](E)\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}\lesssim\|b\|_{\mathrm{BMO}(\mathbb{R}^{n+m})}\|E\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}$$
$$\lesssim\|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|E\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}.$$

As a consequence, from the definition of $[b, P]$ and (5-2) we get

$$\begin{aligned}\left|\int_{\mathbb{R}^{n+m}} E(x,y)\cdot B(x,y)\,b(x,y)\,dx\,dy\right| &= \left|\int_{\mathbb{R}^{n+m}} [b,P]E(x,y)\cdot B(x,y)\,dx\,dy\right| \\ &\lesssim \|b\|_{\mathrm{BMO}(\mathbb{R}^{n+m})}\|E\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}\|B\|_{L^q(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})} \\ &\lesssim \|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|E\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}\|B\|_{L^q(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}.\end{aligned}$$

Thus we get that $E\cdot B$ is in $H^1(\mathbb{R}^{n+m})$ with

$$\|E\cdot B\|_{H^1(\mathbb{R}^{n+m})} \lesssim \|E\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}\|B\|_{L^q(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}.$$

To show the second result, we now define the projection operator $\mathcal{P}^{(2)}$ as

$$\mathcal{P}^{(2)}E = \left(E_{n+1} + R_1^{(2)}\left(\sum_{k=1}^m R_k^{(2)}E_{n+k}\right), \ldots, E_{n+m} + R_{n+m}^{(1)}\left(\sum_{k=1}^m R_k^{(1)}E_{n+k}\right)\right).$$

Then, again, by definition, we have

$$\operatorname{div}_y \mathcal{P}^{(2)}E = 0$$

since

$$\sum_{j=1}^m R_j^{(2)}\left(E_{n+j} + R_j^{(2)}\left(\sum_{k=1}^m R_k^{(2)}E_{n+k}\right)\right) = 0. \tag{5-3}$$

Now fix $x\in\mathbb{R}^n$; by using the definition of $\mathcal{P}^{(2)}$ and the fact (5-3), we get that for $b\in\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$,

$$\int_{\mathbb{R}^m} E(x,y)\cdot_2 B(x,y)b(x,y)\,dy = \int_{\mathbb{R}^m}[b(x,\cdot),\mathcal{P}^{(2)}]E(x,y)\psi(x,y)\,dy.$$

Integrating the above equality over $\mathbb{R}^n$, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^m} E(x,y)\cdot_2 B(x,y)b(x,y)\,dy\,dx\right| &= \left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^m}[b(x,\cdot),\mathcal{P}^{(2)}]E(x,y)\cdot_2 B(x,y)\,dy\,dx\right| \\ &\lesssim \int_{\mathbb{R}^n}\|b(x,\cdot)\|_{\mathrm{BMO}(\mathbb{R}^n)}\|E(x,\cdot)\|_{L^p(\mathbb{R}^m)}\|B(x,\cdot)\|_{L^q(\mathbb{R}^m)}\,dx \\ &\lesssim \|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\int_{\mathbb{R}^n}\|E(x,\cdot)\|_{L^p(\mathbb{R}^m)}\|B(x,\cdot)\|_{L^q(\mathbb{R}^m)}\,dx \\ &\lesssim \|b\|_{\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)}\|E\|_{L^p(\mathbb{R}^m\times\mathbb{R}^n;\mathbb{R}^{n+m})}\|B\|_{L^p(\mathbb{R}^m\times\mathbb{R}^n;\mathbb{R}^{n+m})}.\end{aligned}$$

Here we use again Lemma 4.1 and Hölder's inequality. Taking the supremum over all $b\in\mathrm{bmo}_{\mathcal{F}}(\mathbb{R}^n\times\mathbb{R}^m)$ we obtain that

$$\int_{\mathbb{R}^m}\|E(\cdot,y)\cdot_2 B(\cdot,y)\|_{H^1(\mathbb{R}^m)}\,dy \lesssim \|E\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}\|B\|_{L^q(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^{n+m})}. \qquad \square$$

Acknowledgments

Duong, Li and Pipher are supported by ARC DP 160100153. Duong and Li are also supported by Macquarie University Research Seeding Grant. Wick's research is supported in part by National Science

Foundation grant no. DMS-1560955. This paper started in July 2016 during Li's visit with Pipher at Brown University. Li would like to thank the Department of Mathematics of Brown University for its hospitality. Part of this material is based upon work supported by the National Science Foundation under grant no. DMS-1440140 while Ou and Pipher were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester.

References

[Bényi et al. 2017] A. Bényi, J. M. Martell, K. Moen, E. Stachura, and R. H. Torres, "Boundedness results for commutators with BMO functions via weighted estimates: a comprehensive approach", preprint, 2017. arXiv

[Chang and Fefferman 1980] S.-Y. A. Chang and R. Fefferman, "A continuous version of duality of H^1 with BMO on the bidisc", *Ann. of Math.* (2) **112**:1 (1980), 179–201. MR Zbl

[Chang and Fefferman 1982] S.-Y. A. Chang and R. Fefferman, "The Calderón–Zygmund decomposition on product domains", *Amer. J. Math.* **104**:3 (1982), 455–468. MR Zbl

[Chang and Fefferman 1985] S.-Y. A. Chang and R. Fefferman, "Some recent developments in Fourier analysis and H^p-theory on product domains", *Bull. Amer. Math. Soc.* (*N.S.*) **12**:1 (1985), 1–43. MR Zbl

[Chung et al. 2012] D. Chung, M. C. Pereyra, and C. Perez, "Sharp bounds for general commutators on weighted Lebesgue spaces", *Trans. Amer. Math. Soc.* **364**:3 (2012), 1163–1177. MR Zbl

[Coifman et al. 1976] R. R. Coifman, R. Rochberg, and G. Weiss, "Factorization theorems for Hardy spaces in several variables", *Ann. of Math.* (2) **103**:3 (1976), 611–635. MR Zbl

[Coifman et al. 1993] R. Coifman, P.-L. Lions, Y. Meyer, and S. Semmes, "Compensated compactness and Hardy spaces", *J. Math. Pures Appl.* (9) **72**:3 (1993), 247–286. MR Zbl

[Dalenc and Ou 2016] L. Dalenc and Y. Ou, "Upper bound for multi-parameter iterated commutators", *Publ. Mat.* **60**:1 (2016), 191–220. MR Zbl

[Fefferman 1986] R. Fefferman, "Multiparameter Fourier analysis", pp. 47–130 in *Beijing lectures in harmonic analysis* (Beijing, 1984), edited by E. M. Stein, Ann. of Math. Stud. **112**, Princeton Univ. Press, 1986. MR Zbl

[Fefferman 1987] R. Fefferman, "Harmonic analysis on product spaces", *Ann. of Math.* (2) **126**:1 (1987), 109–130. MR Zbl

[Fefferman 1999] R. A. Fefferman, "Multiparameter Calderón–Zygmund theory", pp. 207–221 in *Harmonic analysis and partial differential equations* (Chicago, IL, 1996), edited by M. Christ et al., Univ. Chicago Press, 1999. MR Zbl

[Fefferman and Stein 1972] C. Fefferman and E. M. Stein, "H^p spaces of several variables", *Acta Math.* **129**:3-4 (1972), 137–193. MR Zbl

[Fefferman and Stein 1982] R. Fefferman and E. M. Stein, "Singular integrals on product spaces", *Adv. in Math.* **45**:2 (1982), 117–143. MR Zbl

[Ferguson and Lacey 2002] S. H. Ferguson and M. T. Lacey, "A characterization of product BMO by commutators", *Acta Math.* **189**:2 (2002), 143–160. MR Zbl

[Ferguson and Sadosky 2000] S. H. Ferguson and C. Sadosky, "Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures", *J. Anal. Math.* **81** (2000), 239–267. MR Zbl

[Głowacki 2010] P. Głowacki, "Composition and L^2-boundedness of flag kernels", *Colloq. Math.* **118**:2 (2010), 581–585. Correction in **120**:2 (2010), 331. MR Zbl

[Gundy and Stein 1979] R. F. Gundy and E. M. Stein, "H^p theory for the poly-disc", *Proc. Nat. Acad. Sci. U.S.A.* **76**:3 (1979), 1026–1029. MR Zbl

[Han and Lu 2008] Y. Han and G. Lu, "Discrete Littlewood–Paley–Stein theory and multi-parameter Hardy spaces associated with flag singular integrals", preprint, 2008. arXiv

[Han et al. 2014] Y. Han, G. Lu, and E. Sawyer, "Flag Hardy spaces and Marcinkiewicz multipliers on the Heisenberg group", *Anal. PDE* **7**:7 (2014), 1465–1534. MR Zbl

[Han et al. 2016a] Y. Han, M.-Y. Lee, J. Li, and B. D. Wick, "Characterizations of flag Hardy space via Riesz transforms, maximal functions and Littlewood–Paley theory", preprint, 2016. arXiv

[Han et al. 2016b] Y. S. Han, C. C. Lin, and X. F. Wu, "Boundedness of singular integrals with flag kernels on weighted flag Hardy spaces", preprint, 2016, available at http://www.ncts.ntu.edu.tw/bib_detail.php?gid=34.

[Hytönen 2016] T. P. Hytönen, "The Holmes–Wick theorem on two-weight bounds for higher order commutators revisited", *Arch. Math.* (*Basel*) **107**:4 (2016), 389–395. MR Zbl

[Journé 1985] J.-L. Journé, "Calderón–Zygmund operators on product spaces", *Rev. Mat. Iberoamericana* **1**:3 (1985), 55–91. MR Zbl

[Kunwar and Ou 2017] I. Kunwar and Y. Ou, "Two-weight inequalities for multilinear commutators", preprint, 2017. arXiv

[Lacey and Terwilleger 2009] M. Lacey and E. Terwilleger, "Hankel operators in several complex variables and product BMO", *Houston J. Math.* **35**:1 (2009), 159–183. MR Zbl

[Lacey et al. 2009] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, "Multiparameter Riesz commutators", *Amer. J. Math.* **131**:3 (2009), 731–769. MR Zbl

[Lacey et al. 2010] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, "Iterated Riesz commutators: a simple proof of boundedness", pp. 171–178 in *Harmonic analysis and partial differential equations* (El Escorial, 2008), edited by P. Cifuentes et al., Contemp. Math. **505**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

[Lacey et al. 2012] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, "Multi-parameter div-curl lemmas", *Bull. Lond. Math. Soc.* **44**:6 (2012), 1123–1131. MR Zbl

[Müller et al. 1995] D. Müller, F. Ricci, and E. M. Stein, "Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups, I", *Invent. Math.* **119**:2 (1995), 199–233. MR Zbl

[Müller et al. 1996] D. Müller, F. Ricci, and E. M. Stein, "Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups, II", *Math. Z.* **221**:2 (1996), 267–291. MR Zbl

[Nagel et al. 2001] A. Nagel, F. Ricci, and E. M. Stein, "Singular integrals with flag kernels and analysis on quadratic CR manifolds", *J. Funct. Anal.* **181**:1 (2001), 29–118. MR Zbl

[Nagel et al. 2012] A. Nagel, F. Ricci, E. Stein, and S. Wainger, "Singular integrals with flag kernels on homogeneous groups, I", *Rev. Mat. Iberoam.* **28**:3 (2012), 631–722. MR Zbl

[Nagel et al. 2018] A. Nagel, F. Ricci, E. M. Stein, and S. Wainger, *Algebras of singular integral operators with kernels controlled by multiple norms*, Mem. Amer. Math. Soc. **1230**, Amer. Math. Soc., Providence, RI, 2018.

[Pipher 1986] J. Pipher, "Journé's covering lemma and its extension to higher dimensions", *Duke Math. J.* **53**:3 (1986), 683–690. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR Zbl

[Wu 2014] X. Wu, "Weighted norm inequalities for flag singular integrals on homogeneous groups", *Taiwanese J. Math.* **18**:2 (2014), 357–369. MR Zbl

Received 19 Feb 2018. Revised 24 Jul 2018. Accepted 16 Sep 2018.

XUAN THINH DUONG: xuan.duong@mq.edu.au
Department of Mathematics, Macquarie University, Sydney, Australia

JI LI: ji.li@mq.edu.au
Department of Mathematics, Macquarie University, Sydney, Australia

YUMENG OU: yumeng.ou@baruch.cuny.edu
Department of Mathematics, Baruch College, The City University of New York, New York, NY, United States

JILL PIPHER: jpipher@math.brown.edu
Department of Mathematics, Brown University, Providence, RI, United States

BRETT D. WICK: wick@math.wustl.edu
Department of Mathematics, Washington University–St. Louis, St. Louis, MO, United States

mathematical sciences publishers

msp

Analysis & PDE
msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2019 is US $310/year for the electronic version, and $520/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 12 No. 5 2019

2157-5045(2019)12:5;1-L